

**ARCHAEA DIVERSITY IN THE COASTAL
WATERS OF WEST PENINSULAR MALAYSIA**

SAW SUET YIT

**DISSERTATION SUBMITTED IN FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE**

**INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2014

ABSTRACT

In this study, we explored the archaea diversity in four locations (Kuala Selangor, Port Klang, Port Dickson, Sungai Muar) along the Straits of Malacca using the cloning method. Four 16S rDNA clone libraries were constructed with archaea specific primers (Arc21F and Arc958R), and in total, 1074 randomly selected clones were screened by PCR-restriction fragment length polymorphism (PCR-RFLP) analysis using three different restriction enzymes (*RsaI*, *CfoI*, and *DdeI*). Majority of the sequences from three clone libraries were affiliated to Euryarchaeota whereas one clone library (Kuala Selangor) was dominated by Thaumarchaeota sequences. Phylogenetic analysis of 16S rDNA sequences showed high archaea diversity with the Shannon index from 2.16 to 3.24. Using the weighted non-normalized principal coordinate analysis, the archaeal communities of the four sites were clearly separated from each other. Via canonical correspondence analysis, the distribution of OTUs could be related to environmental variables especially chlorophyll *a* concentration.

ABSTRAK

Kami mengkaji diversiti arkea di empat lokasi (Kuala Selangor, Port Klang, Port Dickson, Sungai Muar) sepanjang Selat Melaka dengan teknik pengklonan. Empat perpustakaan klon 16S rDNA telah dibina dengan menggunakan primer spesifik untuk deteksifikasi arkea (Arc21F and Arc958R). Sejumlah 1074 klon yang terpilih secara rawak telah disaring dengan teknik analisa PCR-RFLP dengan menggunakan tiga jenis enzim restriksi yang berlainan (*RsaI*, *CfoI*, and *DdeI*). Majoriti sekuen dari tiga perpustakaan klon menunjukkan persamaan dengan filum Euriarkeota, kecuali satu perpustakaan klon (Kuala Selangor) yang menunjukkan dominasi sekuen Thaummarkeota. Analisis filogenetik sekuen 16S rDNA menunjukkan diversiti arkea yang tinggi dengan bacaan indeks Shannon dari 2.16 ke 3.24. Analisa statistik multivariat *Principal Coordinate Analysis* (PCoA) menunjukkan komuniti arkea di empat lokasi yang dikaji ini jelas dipisahkan dari satu sama lain. *Canonical Correspondence Analysis* (CCA) digunakan untuk karakterisasi hubungan distribusi unit operasi taksonomi (OTU) dengan pembolehubah alam sekitar. Analisa CCA menunjukkan distribusi unit operasi taksonomi mungkin dikaitkan dengan pembolehubah alam sekitar, terutamanya kepekatan klorofil *a*.

ACKNOWLEDGEMENT

I would like to extend my deepest appreciation to my supervisor Assoc. Prof. Dr. Lee Choon Weng. This thesis would not be possible without his constant guidance and his essence of genius, which still amazes me where it originated. From this course of study, I have not only gained immense knowledge but also learned the meaning of perseverance and hard work from my supervisor.

I will not forget Dr. Bong Chui Wei's words of advice and moral support, which absorbs the tears I have shed. Your support meant a lot to me. To my lab mates, Ang Pey Pey, Heng Pei Li, Lim Joon Hai and Henry You, thank you for being helpful companions. Work environment would not be lively without you.

To my partner, Wilson Wong, you have always been the understanding one. You have always been there during my every emotional breakdown. Thank you mum and dad, for understanding what I need. Mummy, thank you for the nutritious food that keeps me strong through my journey.

I would also like to thank University of Malaya for providing me with monetary assistance (Fellowship Scheme) enabling me to complete my study. The completion of this thesis means a lot to me.

TABLE OF CONTENTS

Chapter 1: Introduction		
1.0	<i>Archaea, the third domain of life</i>	1
1.1	<i>Molecular approaches in determining diversity and abundance of Archaea</i>	17
1.2	<i>Environmental factors influencing Archaea</i>	23
1.3	<i>Biotechnological aspects of Archaea</i>	25
1.4	<i>Aim of this study</i>	28
Chapter 2: Materials and Methods		
2.0	<i>Sampling</i>	29
2.1	<i>Physico-chemical parameters</i>	33
2.2	<i>Biological parameters</i>	35
2.3	<i>Molecular Methods</i>	36
Chapter 3: Results		
3.0	<i>Physico-chemical characteristics</i>	46
3.1	<i>Molecular results</i>	50
3.2	<i>Archaea diversity and phylogenetic relationship</i>	56
Chapter 4: Discussion		
4.0	<i>Environmental conditions</i>	73
4.1	<i>Uncultured Archaea in different environments</i>	74
4.2	<i>Phylogenetic relatedness</i>	77
4.3	<i>Techniques in evaluating archaeal diversity</i>	79
4.4	<i>The expanding archaeal taxonomy</i>	79
4.5	<i>Ordination analysis of OTUs to environmental variables</i>	80
Chapter 5: Conclusion		84
References		85
Appendix		101

LIST OF FIGURES

Figure 1.0 Phylogenetic tree of the three domains of life (Bacteria, Archaea, and Eucarya) based on the evolutionary distance of the 16S rRNA molecule (adapted from Barns et al., 1996).

Figure 1.1 Schematic tree showing the relationships of the four major groups (Marine Group I, II, III and IV) of planktonic archaea relative to cultivated groups (adapted from DeLong 2003).

Figure 1.2 Phylogenetic analysis of deep-sea benthic archaea (adapted from Vetriani et al., 1999).

Figure 1.3 Maximum likelihood tree showing Thaumarchaeota forming a monophyletic group with hyperthermophilic crenarchaeotas (adapted from Brochier-Armanet et al., 2008).

Figure 1.4 Maximum likelihood phylogenetic tree of concatenated (SSU+LSU) rRNA genes (adapted from Nunoura et al., 2011). Bacterial sequences were used as out-group. Numbers indicate bootstrap values from 100 replications. *Caldiarchaeum subterraneum* has been proposed to be in a new phylum, Aigarchaeota.

Figure 2.0 Map showing the location of the sampling sites at Kuala Selangor (3° 20' 52" N, 101° 15' 6" E), Port Klang (3° 0' 2" N, 101° 23' 29" E), Port Dickson (2° 29' 34" N, 101° 50' 22" E) and Sungai Muar (2° 2' 56" N, 102° 33' 8.9" E).

Figure 2.1 Flowchart of experimental methodology.

Figure 3.0 Gel electrophoresis of extracted plasmids. Lanes 1, 22, 37: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs). Lanes 2-21, 23-36, 38-51: Extracted plasmids.

Figure 3.1 Gel electrophoresis of PCR amplifications of DNA fragments from purified plasmid DNA using Arc 21f/958r primers. Lane 1: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs), Lane 2: Negative control without DNA template, Lane 3-25: Amplicons (~ 900 bp).

Figure 3.2 Comparison of *RsaI* RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

Figure 3.3 Comparison of *CfoI* RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

Figure 3.4 Comparison of *DdeI* RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

Figure 3.5 Rarefaction curve of archaeal clone libraries constructed with MOTHUR using 10,000 iterations. Dotted lines shows the higher and lower confidence intervals of each curve.

Figure 3.6 Archaeal phylotype compositions for 16S rDNA clone libraries constructed from KS, PK, PD and SM.

Figure 3.7 Phylogenetic analysis of Euryarchaeota 16S rDNA sequences obtained from KS, PK, PD and SM. *Cenarchaeum symbiosum* was used as an outgroup. OTUs obtained from this study were bold, and the number after the ` - ` shows the accession code deposited in GenBank. Clone sequences obtained from various environments similar to OTU sequences in this study were shown by stating the isolation source and its accession code. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

Figure 3.8 Phylogenetic analysis of Miscellaneous Crenarchaeotic Group and Thaumarchaeota (SAGMCG-1 and Marine Group 1) 16S rDNA sequences obtained from KS, PK, PD and SM. *Aquifex pyrophilus* was used as an outgroup. OTUs obtained from this study were bold, and the number after the ` - ` shows the accession code deposited in GenBank. Clone sequences obtained from various environments similar to OTU sequences in this study were shown by stating the isolation source and its accession code. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

Figure 3.9 Phylogenetic analysis of archaeal 16S rDNA sequences obtained from KS, PK, PD and SM. *Aquifex pyrophilus* was used as an outgroup. OTUs obtained from this study were bold, and the number after the ` - ` shows the accession code deposited in GenBank. Sequences obtained from various environments similar to OTU sequences in this study were shown by stating the isolation source and its accession code. Italicized words represents isolated cultures. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

Figure 4.0 Diversity of uncultured archaea in studies from environments of different trophic state. Trophic state: 1 = oligotrophic, 2 = mesotrophic, 3 = eutrophic, 4 =

hypertrophic. Trophic state is determined using Carlson's Trophic State Index (Carlson, 1977).

Figure 4.1 Correspondence canonical analysis (CCA) ordination plots for the first two dimensions of CCA of the relationship between the archaeal OTUs with environmental factors. OTUs present only in KS, PD and SM are indicated by points A, B and C, respectively. OTUs present only in PK was indicated by point D.

LIST OF TABLES

Table 1.0 Classification within the phylum Euryarchaeota.

Table 1.1 Molecular techniques utilized in determining Archaea diversity.

Table 1.2 Industrial applications of archaeal products (adapted from Alquéres et al., 2007).

Table 2.0 Sampling dates and description of sampling sites.

Table 2.1 Universal Archaea primer sequences used in this study.

Table 2.2 PCR conditions used in this work. Conditions used for PCR amplification during pre-cloning and plasmid PCR are similar.

Table 3.0 Environmental parameters at KS, PK, PD and SM. Values are means (\pm standard deviations). *, **, *** showed the results of ANOVA testing at $P < 0.05$; $P < 0.01$ and $P < 0.001$, respectively. The same letters of the alphabet are used to indicate values whose means were significantly different.

Table 3.1 Coverage, observed OTUs, richness and diversity indices for clone libraries constructed. Values in parentheses represent lower and higher confidence interval.

Table 3.2 Phylogenetic affiliations of archaeal clone libraries.

LIST OF APPENDIX

Appendix A: Preparation of the reagents used in determination of urea concentration.

Appendix B: Sequences of unique OTU were deposited in the GenBank database under the following accession numbers.

INTRODUCTION

1.0 Archaea, the third domain of life

Archaea is the third domain of life (Figure 1.0) and its evolutionary relationships among archaeal species are established based on rRNA studies. Archaea is once thought to primarily inhabit extreme aquatic and terrestrial environments such as hot springs (Huang et al., 2011), deep-sea hydrothermal vents (Moyer et al., 1998, Edgcomb et al., 2007), hypersaline (Ahmad et al., 2011) and psychrophilic (Margesin and Miteva, 2011) environments but is now commonly found in mesophilic environments based on culture-independent analysis of rRNA gene sequences (Brochier-Armanet et al., 2008, Bergmann et al., 2010).

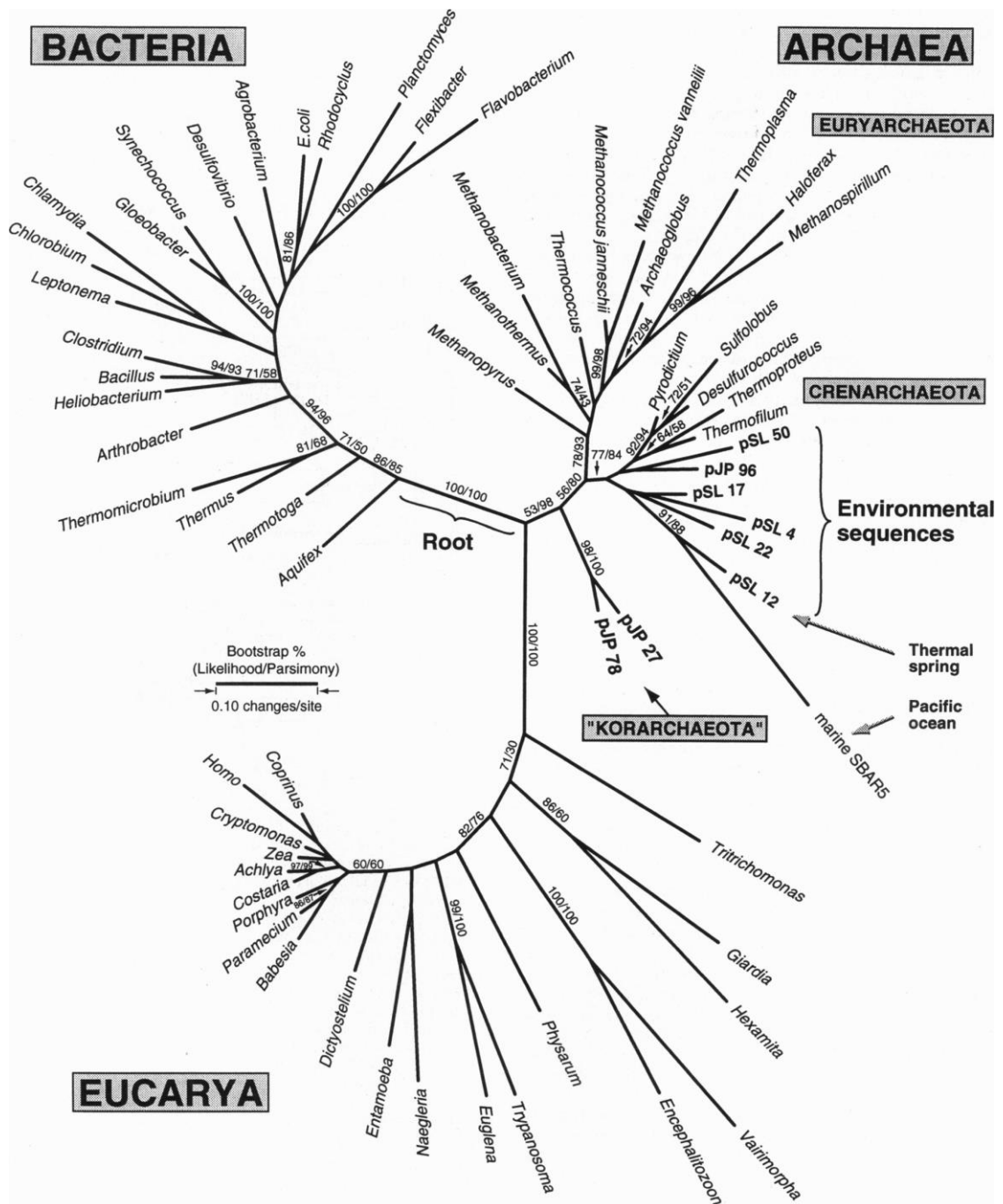


Figure 1.0 Phylogenetic tree of the three domains of life (Bacteria, Archaea, and Eucarya) based on the evolutionary distance of the 16S rRNA molecule (adapted from Barns et al., 1996).

The domain Archaea consists of six main phyla; Euryarchaeota, Crenarchaeota, Thaumarchaeota, Korarchaeota, Nanoarchaeota and Aigarchaeota. Euryarchaeota comprises of physiologically diverse groups; methanogens, extreme halophiles, and some hyperthermophiles (Massana et al. 1997). The phylum Euryarchaeota is further organized into eight classes (Table 1.0): Methanobacteria, Methanococci, Methanomicrobia, Halobacteria, Thermoplasmata, Thermococci, Archaeoglobi and Methanopyri. The methanogens are classified into three distinct classes. Class I methanogens consisted of the orders Methanopyrales, Methanococcales and Methanobacteriales, and the Class II and Class III were composed of the order Methanomicrobiales and the order Methanosarcinales, respectively. Methanogens are strict anaerobes, while extreme halophiles thrive as obligate aerobes in highly saline environment. It is interesting to note that Archaea includes the sole organisms capable of methanogenesis; i.e. production of methane from hydrogen and carbon dioxide or from acetate (Gribaldo and Brochier-Armanet, 2006).

Table 1.0 Classification within the phylum Euryarchaeota.

Phylum: Euryarchaeota		
Class: Archaeoglobi		
Order: Archaeoglobales		
Class: Halobacteria		
Order: Halobacteriales		
Class: Thermococci		
Order: Thermococcales		
Class: Thermoplasmata		
Order: Thermoplasmatales		
Class: Methanobacteria		
Order: Methanobacteriales		Class I Methanogen
Class: Methanococci		
Order: Methanococcales		Class I Methanogen
Class: Methanomicrobia		
Order: Methanomicrobiales		Class II Methanogen
Order: Methanosarcinales		Class III Methanogen
Order: Methanocellales		
Class: Methanopyri		
Order: Methanopyrales		Class I Methanogen

On the other hand, unculturable euryarchaeotal sequences have been retrieved from plankton samples obtained from the Pacific Ocean (DeLong, 1992) which forms a cluster referred to as the Marine Group (MG) – II archaea, which is distantly related to halophiles and methanogens (Figure 1.1).

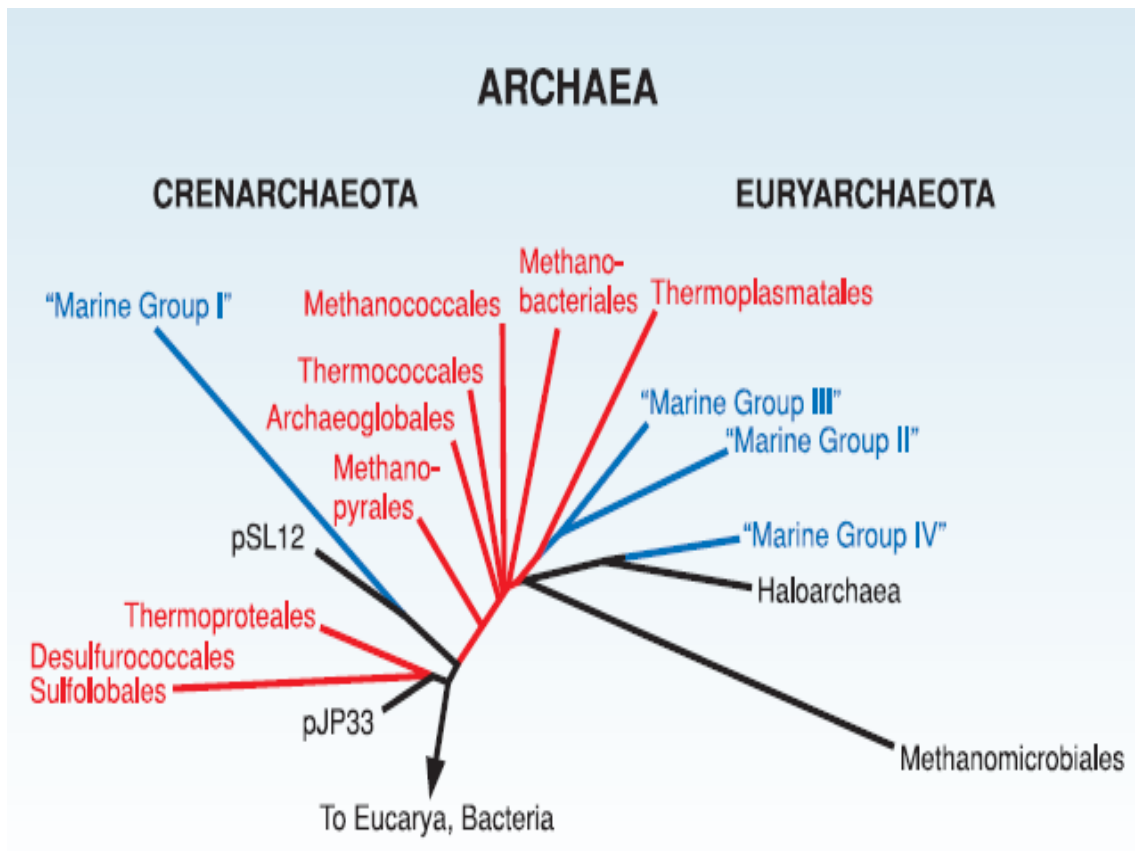


Figure 1.1 Schematic tree showing the relationships of the four major groups (Marine Group I, II, III and IV) of planktonic archaea relative to cultivated groups (adapted from DeLong 2003).

In addition, MG- III, -IV and Marine Benthic Group (MBG) - D and -E are also lineages of euryarchaea, as identified by 16S rDNA sequence. MG-IV represents a group of sequences closely related to haloarchaea; thus suggesting its halophilic characteristic. Being clearly unique from all known planktonic sequences, MG-IV is linked with high depth in the water column, as they are amplified exclusively from deep-sea samples (Lopez-Garcia et al., 2001). MBG-D is equivalent to MG- III, as defined by DeLong (1998). MBG-D is isolated from subsurface marine sediments and has also been observed to be dominant in surface lake waters (Jiang et al., 2008). The phylogenetic position of MBG-E between members of the Methanobacteriales and Methanosarcinales indicates a potential methanogenic phenotype within this group (Vetriani et al., 1999) (Figure 1.2). Similarly, culture-independent molecular analyses has led Takai et al. (2001) to reveal two more groups of euryarchaea; South Africa gold mine Euryarchaeotic Group-1 (SAGMEG-1) and South Africa gold mine Euryarchaeotic Group-2, from the fissure water of South Africa mines. These archaeal phylotypes possess relatively high G + C contents, indicating its thermophilic characteristics. The presence of novel Archaea sequences suggests that these gold mines contain unique thermophilic habitats. Therefore, it is not surprising that many more novel sequences could be retrieved when studying different unexplored habitats with culture-independent methods.

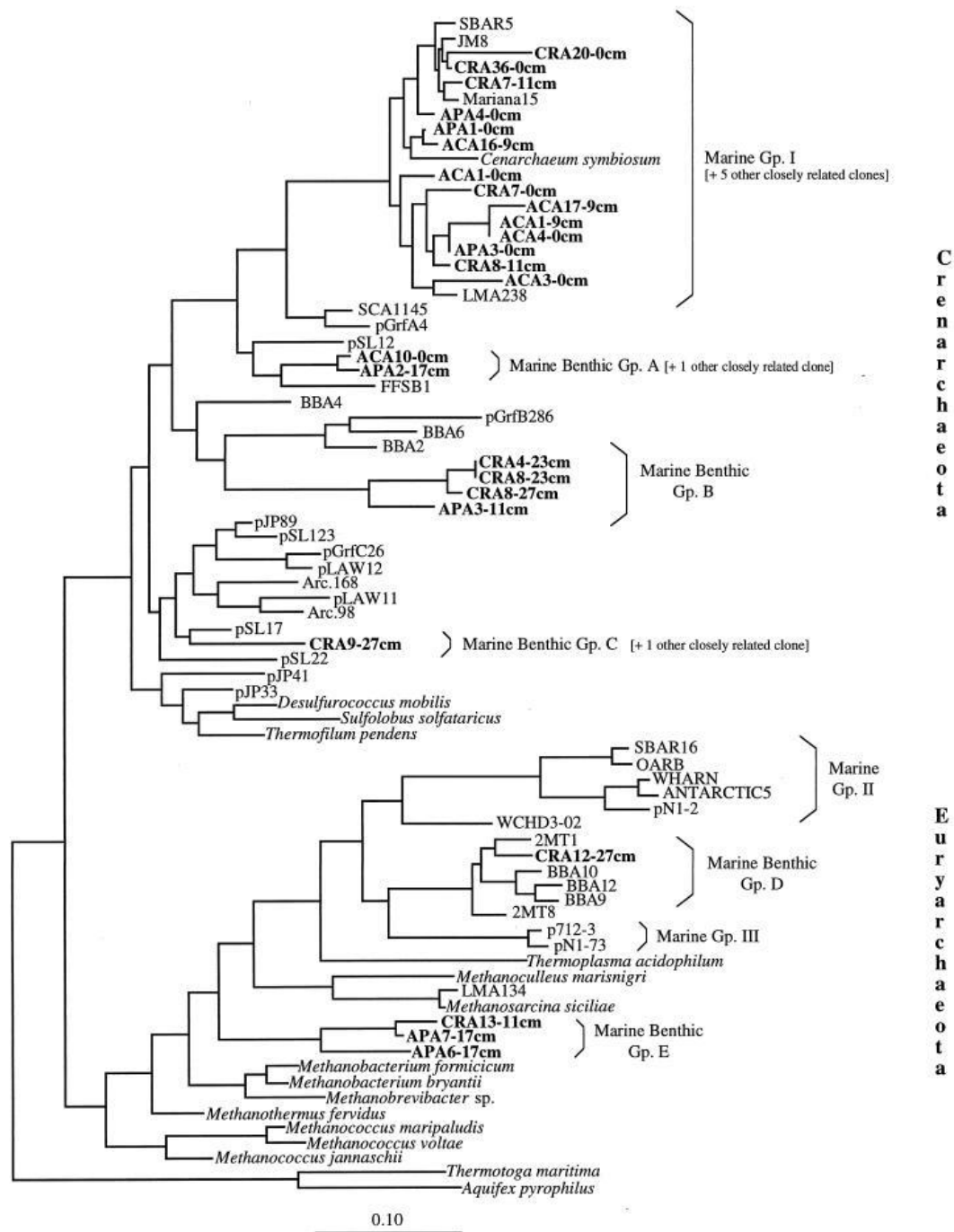


Figure 1.2 Phylogenetic analysis of deep-sea benthic archaea (adapted from Vetriani et al., 1999).

The phylum Crenarchaeota is comprised of five orders: Thermoproteales, Fervidococcales, Acidilobales, Desulfurococcales and Sulfolobales, all within the unique class of Thermoprotei (Boone and Castenholz, 2001). Although most cultivated Crenarchaeota belong to the hyperthermophilic species, crenarchaeotes can be found in both temperature extremes, from boiling water, to freezing water. Various groups of uncultured archaea sequences associated with Crenarchaea have been reported from diverse ecological studies. These include MBG-A, MBG-B and MBG-C (Vetriani et al., 1999), reported from the study of deep-sea sediments. MBG-B Archaea proposed by Vetriani et al. (1999) is synonymous with the Deep-Sea Archaeal Group (DSAG) (Inagaki et al., 2003). The widespread occurrence and metabolically active mesophilic Crenarchaeota in terrestrial and marine environments belongs to the uncultured clade of Miscellaneous Crenarchaeotic Group (MCG). It is notable that the previously designated terrestrial miscellaneous crenarchaeotic group (TMEG) is then changed to MCG after the discovery of marine phylotypes. These mesophilic Crenarchaea forms a deeply divergent clade distantly related to hyperthermophiles. The term “miscellaneous” reflects the extensive habitual range of this group, including subsurface sediments (Zhang et al., 2010), freshwater (Jiang et al., 2008), soil, terrestrial hot springs (Barns et al., 1996) and marine hydrothermal vents (Teske et al., 2002).

The phylum Thaumarchaeota is established based on the first genome sequence of a Group 1 archaeon; the sponge symbiont *Cenarchaeum symbiosum* (Hallam et al., 2006). This recently established phylum is further supported by the genome sequences of the chemolithoautotrophic ammonium oxidizer marine archaeon Candidatus *Nitrosopumilus maritimus*, the first archaea species to be isolated in pure culture (Könneke et al., 2005) and the moderately thermophilic soil

archaeon *Candidatus Nitrososphaera gargensis* (Spang et al., 2010). The first Group 1 sequences are detected in marine environments by DeLong (1992), and are shown to form a distinct lineage from Crenarchaeota. It is vital to keep in mind that the frequently reported Marine Group 1 archaeon, including marine Group 1.1a and soil Group 1.1b from studies before the establishment of the phylum Thaumarchaeota, are at present being classified under the Thaumarchaeota. These mesophilic thaumarchaeotes form monophyletic groups (bootstrap value of 99%) with hyperthermophilic crenarchaeotes (Figure 1.3).

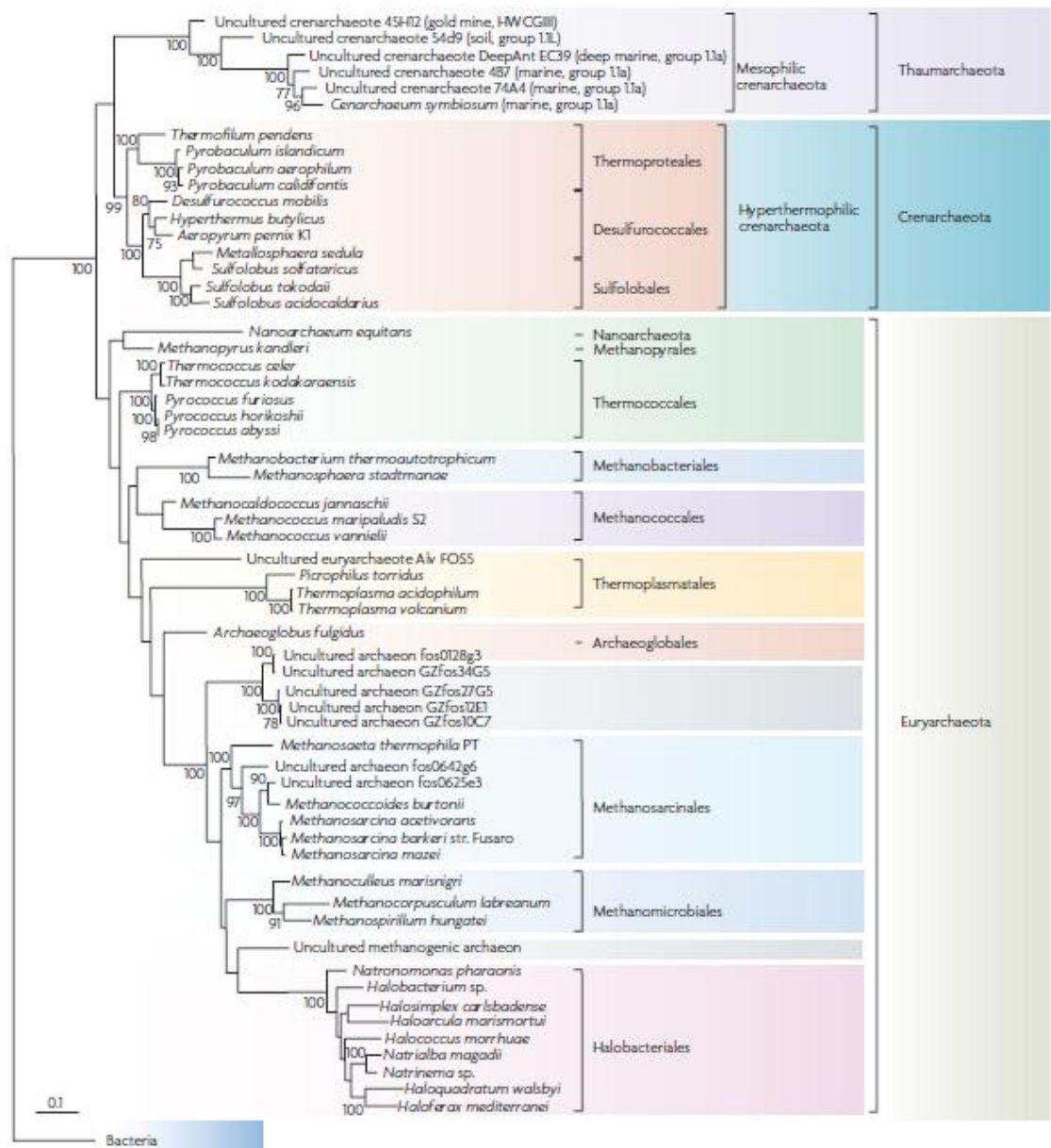


Figure 1.3 Maximum likelihood tree showing Thaumarchaeota forming a monophyletic group with hyperthermophilic crenarchaeota (adapted from Brochier-Armanet et al., 2008).

A large number of uncultivated archaea 16S rRNA sequences recovered from environmental studies shows affiliation to the Thaumarchaeota. For instance, members of group 1.1c (restricted to acidic soils), SAGMCG-1 (subsurface mine), ALOHA group (open ocean), pSL12 group (hot spring) and the hot water crenarchaeotic group (HWCG)- III / *Nitrosocaldus* group (hot springs/ hydrothermal vents) (Pester et al., 2011). Although this novel phylum comprises all known archaeal ammonia oxidizers, it is possible that not all members of this phylum are capable of ammonia oxidation. The energy metabolism of several clusters of environmental sequences within this phylum is yet to be determined. It has been shown that the membrane lipid crenarchaeol, which has been appropriately renamed to thaumarchaeol (Pester et al., 2011), is present in all analyzed ammonia-oxidizing archaea (AOA). Nevertheless, further in-depth studies of uncultivated thaumarchaeons are necessary to verify that thaumarchaeol represents a signature lipid for members of Thaumarchaeota. The presence of Thaumarchaeota in diverse environments, ranging from mesophilic to thermophilic environments acknowledges its physiological versatility. Dominance of Thaumarchaeota is observed in Tibetan hot springs (Huang et al., 2011) and the thermophilic group 1.1b Thaumarchaeote *N. gargensis* is reported to adapt to low ammonium concentrations (Hatzenpichler et al., 2008), indicating a widespread distribution of oligotrophic ammonia oxidizing members of the phylum Thaumarchaeota.

The presence of the phylum Korarchaeota is determined only by environmental DNA sequences (Barns et al., 1996). Korarchaeota are thermophilic organisms which are among the most primitive of all life forms. This phylum includes sequences from hyperthermophilic environments and exclusively uncultivated species, as well as the ancient archaeal group (AAG) (Takai and

Horikoshi, 1999). The first korarchaeal genome being physically isolated from an enrichment culture inoculated with sediments from Obsidian Pool, Yellowstone belonged to Candidatus *Korarchaeum cryptofilum* (Elkins et al., 2008). Genome analysis suggests Candidatus *Korarchaeum cryptofilum* possess a physiology based on peptide fermentation. In addition, the harboring of crenarchaeal and euryarchaeal-like genes support the deep-branching position of Korarchaeota in the archaeal lineage. Cultivation-independent census of Korarchaeota has revealed its presence, although in low abundance, in both terrestrial and marine habitats. For instance, Auchtung et al. (2011), provides evidence of an indigenous community of Korarchaeota in hot springs of Kamchatka, Russia. Screening of deep-sea hydrothermal vent niches at the East Pacific Rise (Auchtung et al., 2006) with Korarchaeota-specific primers identified additional Korarchaeota phylotypes. Reigstad et al. (2010) determines the diversity, distribution and abundance of Korarchaeota by analyzing 19 terrestrial hot springs from Iceland and Kamchatka, Russia, revealing the minority of Korarchaeota in Korarchaeota-positive hot springs. Korarchaeota are not detected in a variety of cooler temperature settings. The high G+C content of its rRNA may explain its thermophilic preference (Dalgaard and Garrett, 1993). Results of Miller-Coleman (2012) expand the geographical and geochemical range of members of the Korarchaeota and further acknowledge the low phylogenetic diversity and endemism of Korarchaeota.

Careful microscopy inspection of the crenarchaea *Ignicoccus* obtained from hot rocks taken from the hydrothermal system at Kolbeinsey ridge, north of Iceland led to the finding of tiny coccoid cells closely attached on the surfaces of *Ignicoccus* cells. With a cell diameter of only 400 nm, these coccoidal cells may be found occurring singly, in pairs or more than 10 cells on its symbiont, *Ignicoccus*. The

unique sequences of this coccoid cell, harboring many base exchanges in the highly conserved regions led to the proposed phylum Nanoarchaeota (the dwarf archaea), and the corresponding species *Nanoarchaeum equitans* (riding the sphere) which portrays the symbiotic relationship, living as a symbiont, possibly as parasites of the crenarchaeote *Ignicoccus*. This association is the first known example of a parasitic / symbiotic partnership involving two archaea, and moreover two hyperthermophilic organisms. Although able to occur on its own, *Nanoarchaeum* can only replicate when attached to its symbiont, *Ignicoccus*. Nanoarchaeota phylum (Huber et al., 2002) harbors the smallest genomes (500 kilobases) of all known prokaryotes. With a highly reduced genome, the nano-sized *Nanoarchaeum equitans* has virtually no obvious metabolic or energetic capabilities and, using unknown mechanisms, must obtain metabolites and energy from *Ignicoccus hospitalis* by attaching to its surface (Huber et al., 2002). To facilitate detection of Nanoarchaeota by fluorescence in situ hybridization (FISH), new oligonucleotide probes has been redesigned based on Nanoarchaeota sequence as 16S rRNA-targeted oligonucleotide probes directed against Crenarchaeota and Euryarchaeota (for example, EURY498R, CREN499R) (Burggraf et al., 1994) failed to stain cells of *Nanoarchaeum equitans* (Huber et al., 2002). At present, the phylum Nanoarchaeota harbors one genus with one species: *Nanoarchaeum equitans* (Huber et al., 2002). A study by Hohn et al. (2002) revealed the presence of Nanoarchaeota 16S rDNA sequences in hydrothermal biotopes in the deep sea, shallow marine areas and solfataric fields located on different continents. Subsequently, results of Casaneuva et al. (2008) shows that nanoarchaeotes are not obligate hyperthermophiles, with the findings of novel archaeal phylotypes in mesophilic hypersaline water and sediment samples.

Analysis of a composite genome sequence of the first representative of the uncultured lineage Hot Water Crenarchaeotic Group I (HWC GI), from the metagenomic library has uncovered unique genomic characteristics that are distinct from earlier reported archaeal genomes. DNA isolation of Candidatus '*Caldiarchaeum subterraneum*' is performed from a sample obtained from a microbial mat community in a geothermal water stream, in which HWC GI dominates. An intrinsic trait of Eucarya; the eukaryote-type protein modifier system was observed in Candidatus '*Caldiarchaeum subterraneum*'. In addition, the genome of *C. subterraneum* reveals the presence of a type I DNA topoisomerase IB (TopoIB) family that has been found only in the Thaumarchaeota, and lost in the Euryarchaeota and hyperthermophilic Crenarchaeota. Phylogenetic analysis based on concatenated SSU+LSU rRNA gene (Figure 1.4), concatenated ribosomal proteins and RNA polymerase subunits and translational elongation factor 2 all expressed that *C. subterraneum* forms a robust cluster with Thaumarchaeota, and is distinct from the hyperthermophilic Crenarchaeota. Although so, the genomes of Thaumarchaeota present more euryarchaeotic features and *C. subterraneum* conserves more crenarchaeotic features. This led to the proposal of a novel archaeal phylum, tentatively called 'Aigarchaeota'. However, more genomic studies is required to show if 'Aigarchaeota' represents a new archaeal phylum or will be classified as deep-branching member of the Crenarchaeota or Thaumarchaeota (Nunoura et al., 2011). The expanding collection of environmental sequences have contributed to our further understanding of archaeal diversity and at the same time, shows that the archaeal phylogenetic tree is more complicated than expected.

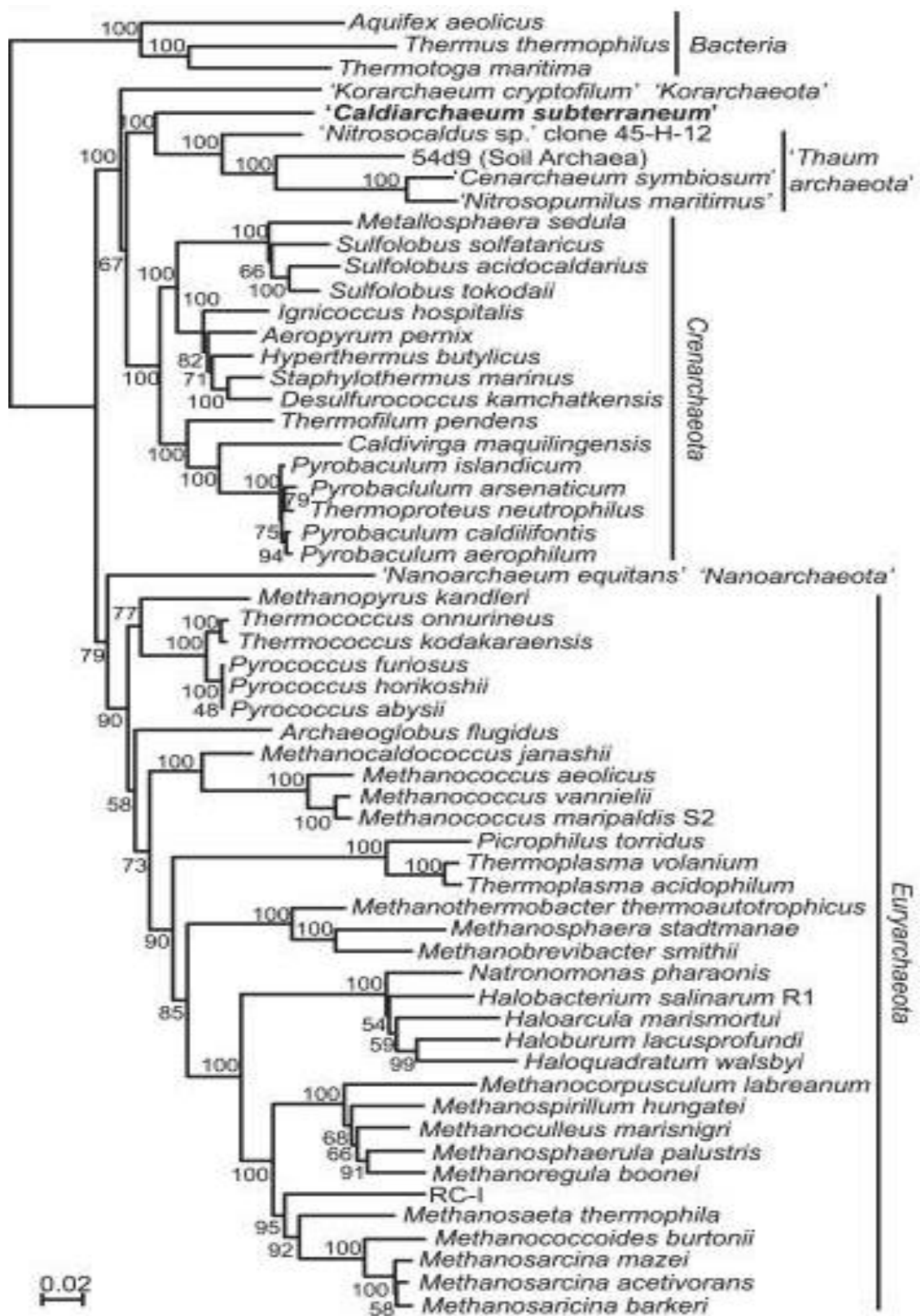


Figure 1.4 Maximum likelihood phylogenetic tree of concatenated (SSU+LSU) rRNA genes (adapted from Nunoura et al., 2011). Bacterial sequences were used as out-group. Numbers indicate bootstrap values from 100 replications. *Caldiarchaeum subterraneum* has been proposed to be in a new phylum, Aigarchaeota.

1.1 Molecular approaches in determining diversity and abundance of Archaea

The abundance and composition of archaea is important in the study of environmental microbiology. The ubiquity and abundance of mesophilic archaea has been proven with the vast microbial ecology studies carried out in a wide variety of habitats. Terrestrial environments (Buckley et al., 1998) such as mangrove soil (Yan et al., 2006) and barley field (Poplawski et al., 2007) and aquatic ecosystems, such as marine water columns and sediments (Francis et al., 2005), mucus of corals (Siboni et al., 2008), deep marine sediments (Inagaki et al., 2006), and oceans (Wuchter et al., 2006, Lam et al., 2007, Mincer et al., 2007) do harbour archaea.

DeLong (1992) reported evidence for the widespread occurrence of archaea in oxygenated coastal surface waters and suggested that high archaeal cell densities were a common feature of the world's oceans. In addition, phylogenetic analyses of planktonic archaea indicate that marine archaea generally group within the Crenarchaeota and Euryarchaeota lineages (Church et al., 2003) and these two major phyla can account for as much as one-third of all prokaryotic cells in the global oceans (Karner et al., 2001). The wide distribution of archaea in oxic coastal surface waters shows that these microorganisms represent undescribed physiological types of archaea, which reside and compete with aerobic, mesophilic eubacteria in marine coastal environment (DeLong, 1992). Interestingly, uncultured mesophilic archaea are as ubiquitous in fermented seafood as in terrestrial and aquatic niches (Roh et al., 2010).

The vital roles of microorganisms in nutrient cycling and food webs in the marine and estuarine environment increase the desire for knowledge of diversity of

microorganisms in microbial communities to better understand the complexity of the marine and estuarine environment. Modern genomic approaches have allowed scientists to gain further insights on microbes which have resisted cultivation.

Konneke et al. (2005) succeeded in isolating a marine crenarchaeote that grows chemolithoautotrophically by aerobically oxidizing ammonia to nitrite, leading to the first observation of nitrification in Archaea. Goh et al. (2006) isolated halophilic archaea on a medium which mimics hypersaline environment. Although some archaea have been cultivated, most archaea remains unculturable. However, many new tools have been developed to facilitate the study of the abundance and the diversity of these unculturable archaea.

In order to gain new insight into the archaeal communities which resist cultivation, can be studied with the 16S rRNA gene clone library approach. Although this method enables the study of archaeal community, the analysis of clones is time consuming. Via this approach, Yan et al. (2006) constructed an archaeal 16S rRNA gene library from mangrove soil, and concluded that majority of the archaeal members of mangrove soil were marine in origin.

Fingerprinting techniques are widely available recently. Denaturing gradient gel electrophoresis (DGGE), temperature gradient gel electrophoresis (TGGE), PCR-restriction fragment length polymorphisms (RFLP) and terminal- restriction fragment length polymorphisms (T-RFLP) are routinely used in both terrestrial and aquatic samples for investigating spatio-temporal dynamics of archaea diversity (Table 1.1). DGGE provides both rapid comparison data for many communities and specific phylogenetic information can be obtained from sequencing of excised bands.

However, different phylotypes with very similar electrophoretic mobilities may reduce the quality of the sequences when an excised band is directly sequenced after PCR amplification (Diez et al., 2001). Although laborious, cloning of excised bands may produce much cleaner sequences when complex diversity is involved. The pursuit to characterize microbial communities has now reached a new phase with the growth of next-generation sequencing techniques (NGS), leading towards a more comprehensive description of the microbial community than cloning and sequencing. Zinger et al. (2012) discusses capacities and limitations of the available methodologies in gauging microbial diversity in aquatic ecosystems. Table 1.1 shows the various molecular techniques utilized in determining archaeal diversity in various environmental sources.

The ability to determine the relative abundance of Archaea to Eubacteria is made possible with the development of improved quantification methods. Single-cell fluorescent in situ hybridization (FISH) with rRNA-targeted probes provides a method of directly detecting and quantifying whole, intact, individual picoplankton cells, enabling direct microscopic enumeration using an epifluorescence microscope. The abundance of planktonic Archaea and Bacteria in various aquatic ecosystems has been studied using this technique (DeLong et al., 1999).

Intensification of the fluorescent signal and sensitivity of FISH technique have also been improved with the Catalyzed Reporter Deposition- Fluorescent In-Situ Hybridization (CARD-FISH) technique. This technique employs an in-situ amplification method based on the deposition of a large number of labeled tyramine molecules by activity of a horseradish peroxidase which is coupled to a chosen oligonucleotide probe of particular specificity. Stronger fluorescent signal intensity

of hybridized cells is useful for detecting small and slower growing microorganisms that are otherwise difficult to detect. CARD-FISH can be employed in the quantification of 16S rRNA gene, as well as functional genes (Herfort et al., 2009).

Real-time quantitative PCR (QPCR) has also been widely utilized in the quantification of 16S rRNA gene of archaea in samples from various environments such as lakes (Llirós et al., 2010), boreal mires (Juottonen et al., 2008), seafloor basal (Einen et al., 2008) and many more.

Table 1.1 Molecular techniques utilized in determining Archaea diversity.

Molecular Method	Archaea source	References
PCR-RFLP	Biogas plant	Bergmann et al. (2010)
	Estuarine region of East China Sea	Zeng et al. (2007)
	Mangrove soil	Yan et al. (2006)
	Tropical corals	Kellogg (2004)
	Pine mycorrhizospheres	Bomberg et al. (2003)
	Marine	Massana et al. (1997)
T-RFLP	Sea ice	Collins et al. (2010)
	Coastal Black Sea	Stoica (2009)
	Deep-sea sediments	Luna et al. (2009)
	Marine sediments	Braker et al. (2001)
	Deep South African gold mines	Takai et al. (2001)
DGGE	Epipelagic waters	Lliros et al. (2010)
	Rivers	Herfort et al. (2009)
	Radioactive thermal spring	Weidler et al. (2008)
	Terrestrial hot springs	Perevalova et al. (2008)
	Hypersaline stratified lake	Cytryn et al. (2000)
TGGE	Wastewater treatment plants	Gomez-Silvan et al. (2010)
Pyrosequencing	Soil	Leininger et al. (2006)
	Arctic ocean	Galand et al. (2009)
	Fermented seafood	Roh et al. (2010)

1.2 **Environmental factors influencing Archaea**

The ocean covers 71% of earth's surface of which 40% are tropical oceans (Longhurst and Pauly, 1987). Tropical waters are vast in area and diverse in ecological conditions, but previous studies on archaea have been focused so far on extreme environments, and temperate waters. In Auguet et al.'s (2010) study of global ecological patterns in uncultured archaea, only 3 out of 67 environments studied are from tropical marine waters. The archaeoplankton communities of Guanabara Bay, which represents a tropical impacted estuarine is described by Vieira et al. (2007). Archaeal phylotypes from sediments of the tropical Western Pacific has indicated that similar phylotypes are not restricted to a particular environment (Zhang et al., 2010). Although there are some studies in tropical aquatic ecosystems (Vieira et al., 2007, Zhang et al., 2010) the archaeal community in Strait of Malacca remains to be revealed.

Archaea are versatile in terms of adaptations, to a wide range of environment. Methanogens which exhibits a strictly anaerobic metabolism has been recovered from oxygenated water layers (Lliros et al., 2010). Martens-Habbena et al., (2009) demonstrated that “*Candidatus Nitrosopumilus maritimus*” strain SCM1 is adapted to extreme nutrient limitation. According to these authors, *Nitrosopumilus*-like ammonia oxidizing archaea (AOA) may benefit from this adaptation to compete for nitrogen sources. Even members of *Halobacteriales*, which are often termed as halophilic, are found to exist in abundance in high numbers in low salt condition of a spring (Elshahed et al., 2004).

Although versatile, archaeal populations are influenced by environmental factors; one of which is seasonality, as observed in the polar marine environments,

where there is a decline of the number of archaea from winter to summer (Murray et al., 1998). Within the archaeal community, Crenarchaeota showed dominance from autumn to early spring, while Euryarchaeota were more abundant in summer and early autumn (Wuchter et al., 2006). Archaea has also shown dominance in a particular microbial community with environmental conditions such as high pH and high temperature (Robertson et al., 2005). In addition, Stoica et al. (2007) observed significant correlations between ammonia, nitrate and crenarchaeotal abundance, and salinity played a significant role in the archaeal distribution pattern when phylogenetic patterns and environmental factors were analysed using archaeal 16S rRNA gene sequences (Auguet et al., 2010). Marine archaea exhibits spatial variation where they are generally more abundant below the photic zone (Fuhrman and Ouverney, 1998). On the other hand, biological factors such as total phytoplankton biomass and also its community composition have been shown to influence crenarchaeotal abundance in the North Sea, making these factors, together with nutrient concentrations useful for predicting crenarchaeotal abundance (Herfort et al., 2007). There was no specified presence of different types of archaeal groups in terms of free-living or particle fraction. However, archaeal communities were remarkably different in riverine, coastal and marine waters (Galand et al., 2008).

1.3 Biotechnological aspects of Archaea

As archaea is capable of thriving in extreme environments, there are intense efforts to unveil potential biotechnological applications of their stable cellular components. Many useful enzymes have been isolated from these extremophiles through cloning of genes encoding these enzymes into mesophilic host cells. Most have industrial applications as summarized in Table 1.2 (Alquéres et al., 2007). Although archaea has been utilized in numerous applications, there are still large reservoirs of genes which are of great interest in biotechnology. The main requirement is the availability of pure, intact, high molecular weight DNA. Although unculturable, the genome fragments are still useful for expressing protein-encoding genes. This gives another reason to as why the knowledge of the diversity of uncultivated archaea is important.

Table 1.2 Industrial applications of archaeal products (adapted from Alquéres et al., 2007).

Phenotype	Condition	Product	Application
Thermophile	High temperature (45 - 110°C)	Amylases	Glucose, fructose for sweeteners
		Xylanases	Paper bleaching
		Proteases	Baking, brewing, detergents
		DNA polymerases	Genetic engineering
Psychrophile	Low temperature (>15°C)	Proteases	Dairy production
		Dehydrogenases	Biosensors
		Amylases	Polymer degradation in detergents
Acidophile	Low pH (0 - 4)	Sulfur oxidation	Desulfurization of coal
Alkalophile	High pH (8 - 11)	Cellulases	Polymer degradation in detergents
Halophile	High salt concentration	Whole microorganism	Biopolymers
Piezophile	High pressure	Whole microorganism	Formation of gels and starch granules

Metalophile	High metal concentration	Whole microorganism	Bioremediation, metal biomineralization
Radiophile	High radiation levels	Whole microorganism	Bioremediation of radionuclide contaminated sites

1.4 Aim of this study

Little is known about biogeography and ecological characteristics of archaeal diversity especially in tropical coasts, rivers, and lakes (Vieira et al., 2007). The desire for new environmental archaeal 16S rRNA sequences, including recognition of their phylogenetic affiliation to specific archaeal lineages, is crucial for the further understanding of the environmental preferences of these uncultured archaea and to obtain a clearer picture of archaeal diversity and phylogeny (Auguet et al., 2010). This study aims to give a better insight of phylogenetic diversity and spatial distribution of archaeal communities in four stations along the west coast of Peninsular Malaysia, namely Port Dickson, Port Klang, Kuala Selangor and Sungai Muar. Physico-chemical parameters were also measured in these sites to show the correlation of environmental physico-chemical parameters with the existence and diversity of Archaea.

MATERIALS AND METHODS

2.0 Sampling

Seawater samples were collected about 0.1 m depth at four sampling stations i.e. Kuala Selangor (KS) (3° 20' 52" N, 101° 15' 6" E), Port Klang (PK) (3° 0' 2" N, 101° 23' 29" E), Port Dickson (PD) (2° 29' 34" N, 101° 50' 22" E) and Sungai Muar (SM) (2° 2' 56" N, 102° 33' 8.9" E). These stations are situated along the west coast of Peninsular Malaysia, and face the Straits of Malacca (Figure 2.0). Straits of Malacca is one of the busiest shipping lanes in the world (Yap et al., 2002). KS, PK and SM are mangrove-lined estuaries whereas PD is a sandy coastal system, and each station was sampled three times (Table 2.0). Physico-chemical and selected variables were measured in each sampling. However the clone library was constructed once from each station.

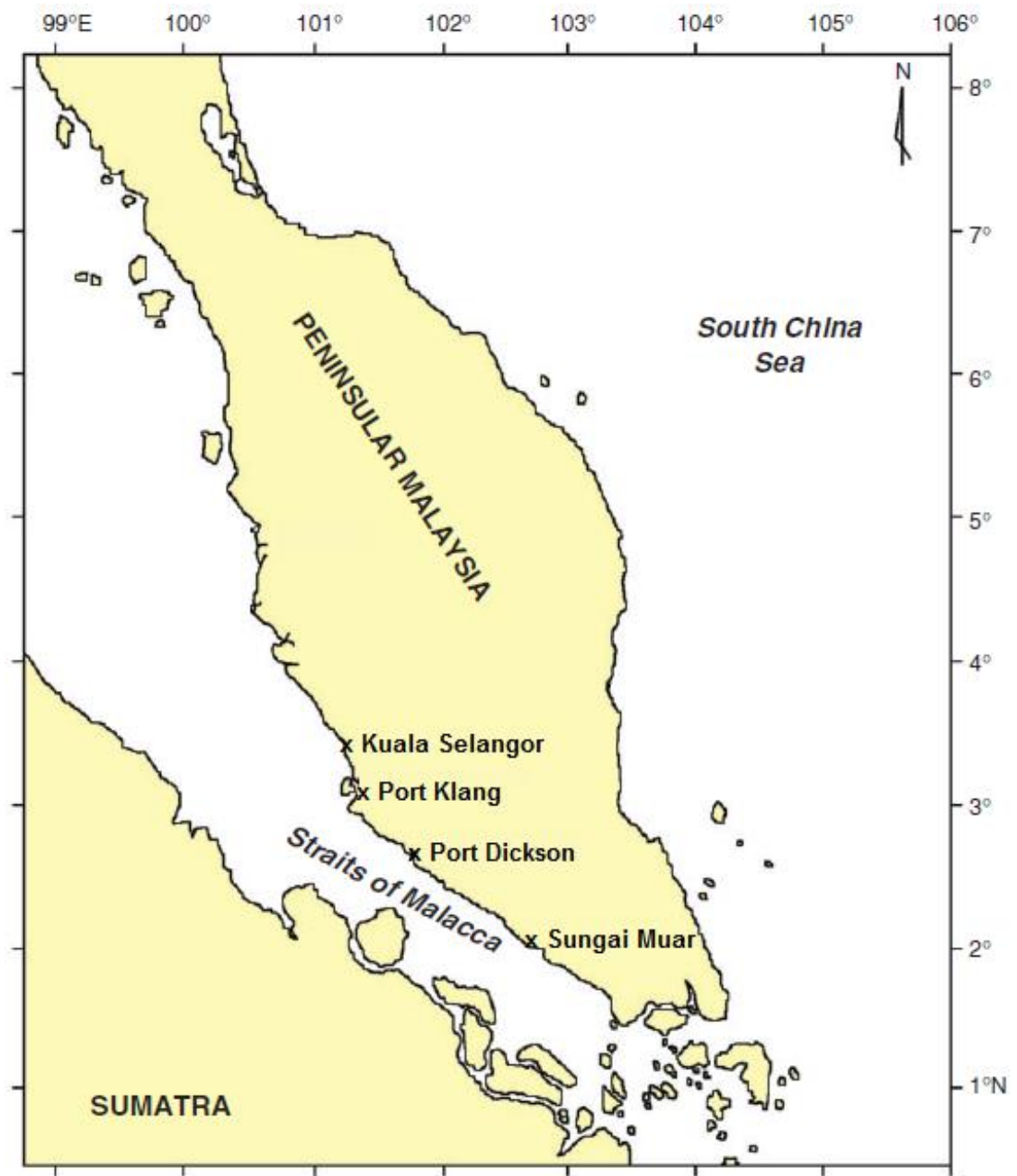


Figure 2.0 Map showing the location of the sampling sites at Kuala Selangor (3° 20' 52" N, 101° 15' 6" E), Port Klang (3° 0' 2" N, 101° 23' 29" E), Port Dickson (2° 29' 34" N, 101° 50' 22" E) and Sungai Muar (2° 2' 56" N, 102° 33' 8.9" E).

Table 2.0 Sampling dates and description of sampling sites.

Sampling sites	Sampling date	Site description	Reference
Kuala Selangor (<i>n</i> = 3)	1 Dec 2010	<ul style="list-style-type: none">• Kuala Selangor is a fishing village situated at the river mouth of Sungai Selangor.	http://www.climateavenue.com/forest.mangr.ecotourism.htm
	27 Dec 2010		
	22 Feb 2011	<ul style="list-style-type: none">• Sungai Selangor is one of the major river systems in Selangor which drains into the Straits of Malacca.• This river is an important source of water supply for domestic and agriculture use, fishing industries for those living along the riverbanks and it supports ecotourism.	
Port Klang (<i>n</i> = 3)	26 Oct 2010	<ul style="list-style-type: none">• Port Klang is Malaysia's principal gateway and busiest port.	http://www.pka.gov.my/phocadownload/gateway/gateway%202011.pdf
	23 Nov 2010		
	23 Dec 2010	<ul style="list-style-type: none">• Surrounding islands form a natural enclosure and make Port Klang well sheltered.	
Port Dickson (<i>n</i> = 3)	22 Oct 2010	<ul style="list-style-type: none">• Port Dickson is a well-known recreational area in Malaysia.	Law et al. (2002)
	9 Nov 2010		
	8 Dec 2011	<ul style="list-style-type: none">• Beach resorts are built along the sampling site. Port Dickson are frequently visited by people and the water quality and microbial community may vary from waters	

of other kinds.

- Impacts of tourists, shipping, oil tankers, refineries, land reclaiming activities, coastal zone management construction and the insufficient sewage water treatment are contributing factors towards deterioration of water quality in Port Dickson.

Sungai Muar (<i>n</i> = 3)	15 Dec 2010	<ul style="list-style-type: none"> • Sungai Muar is a river which flows through the states of Negeri Sembilan and Johor into the Straits of Malacca. This river also flows through Muar town. Tanjung Emas is where the river joins the sea. • Nearby factories which discharge effluents into Sungai Muar may be one of the contributing factors to its deteriorating water quality. • Sungai Muar has also become narrower and shallower over the years, contributing to the occurrence of floods in the nearby areas, classified as a flood-prone area by the Department of Irrigation and Drainage Malaysia. 	Hashim (2010)
	9 Feb 2011		
	6 May 2011		

2.1 Physico-chemical parameters

Surface water temperature, salinity, pH and Secchi disc depth of the sampling sites were measured in-situ. Surface water temperature was measured with a portable thermometer (Comark PDT 300, Korea). Salinity was measured using a salinometer (YSI 30-50, USA) whereas pH was measured with a portable pH meter (Martini, Mi 106, Romania). For dissolved oxygen concentration, the sample was fixed immediately with manganous chloride and alkaline iodide solutions in site, and determined by the Winkler titration method (Grasshoff et al., 1999) later in the lab. Samples were kept in a cooler box until processing within four hours. One sample was also preserved with glutaraldehyde (1% final concentration) for the determination of bacterial abundance.

In the laboratory, seawater samples were filtered through pre-combusted (500°C for 3 hours) Whatman GF/F filters and stored at -20°C until dissolved inorganic nutrient analysis. The filters were also used for chlorophyll *a* (Chl *a*) and total suspended solids (TSS) determination. For TSS, it was determined by the weight increase after drying the filter (50°C for 72 hours). The dried filter was later ashed in a furnace (500°C for 3 hours) and the weight loss on ignition was calculated as particulate organic matter (POM).

Dissolved inorganic nutrients (ammonium, nitrate, nitrite, phosphate and silicate) were measured according to Parsons et al. (1984) using a spectrophotometer (Hitachi U-1900, Japan). Ammonium was determined based upon the oxidation reaction with hypochlorite in an alkaline medium. Formation of blue indophenol colour was measured at 640 nm. Prior to analysis, strength of hypochlorite used was

tested. Nutrient bottles were incubated in dark as high light intensities may cause overdevelopment of the blue colour.

Nitrate was first reduced by granulated copper-cadmium before being measured as nitrite. As nitrate is reduced to nitrite, the sum of nitrate and nitrite was measured. This reading was then compared with nitrite concentration. For nitrite measurement, sulfanilamide in an acid solution was used to react with the nitrite in the seawater sample. The resulting diazo compound was then reacted with N-(1-naphthyl)-ethylenediamine. Intensity of the coloured azo dye formed was measured with a spectrophotometer at 543 nm.

Concentration of phosphate in seawater sample was determined based on the reaction with a composite reagent containing ascorbic acid, trivalent antimony and molybdic acid. A complex was formed and reduced to give a blue solution. Absorbance was measured immediately at 880 nm. Silicate in seawater sample was allowed to react with molybdate. A reducing solution was then added which reduces the silicomolybdate complex formed to give a blue colour. Absorbance was read at 810 nm. Silicate test was carried out in plastic bottles.

In this study, urea concentration was also measured (Goeyens et al., 1998). Seawater sample was collected with an acid washed 5.5 litre bottle, and immediately processed upon returning to the laboratory. Seawater samples were filtered through pre-rinsed GF/F filters. Reagent A (diacetylmonoxime and thiosemicarbazide) and reagent B (sulphuric acid and ferric chloride) (Appendix A) were then added separately to the sample in a dark bottle. The sample was incubated at $(25^{\circ}\text{C} \pm 1)$ in

the dark for 72 hours. The absorbance was measured at 520 nm, and compared against reagent blanks prepared using ultrapure water.

All nutrient measurements above were carried out in triplicates. Coefficient of variation (CV) were < 5% for NH_4^+ , NO_2^- , SiO_4^{4-} , PO_4^{3-} and urea analyses and < 10% for NO_3^- analysis.

2.2 Biological parameters

Chl *a* was measured as a proxy for primary producers, and their pigments were extracted by adding 8 ml of 90% acetone into a dark vial together with the GF/F filter and stored at -20°C overnight. After incubation, sample was centrifuged at 3,000 rpm for 5 minutes, and the Chl *a* concentration was estimated with a spectrophotometer at 750 nm, 665 nm, and 664 nm (Parsons et al., 1984). An average reading from three filters was used for calculation, and the coefficient of variation was < 15%.

For the determination of bacterial abundance, samples were stained with 4'6-diamidino-2-phenylindole (DAPI) ($0.1\ \mu\text{gL}^{-1}$ final concentration) for 7 minutes and slides for direct count were prepared according to Porter and Feig (1980). Prepared filter was examined under an epifluorescence microscope (Olympus BX60, Japan) with a U-MWU filter cassette (excitation, 330 to 385 nm; dichroic mirror, 400 nm; barrier, 420 nm). More than 300 cells or a minimum of thirty fields were counted for each slide. Correction for phototrophic picoplankton was carried out by observing for autofluorescence with a U-MWG filter cassette (excitation, 510 to 550 nm; dichroic mirror, 570 nm; barrier, 590 nm). DAPI staining was done in triplicates, and coefficient of variation was < 30%.

2.3 Molecular Methods

DNA Extraction

Upon reaching the laboratory, seawater sample was filtered onto 0.2 μm pore size filter. All filter papers were kept frozen in -20°C until DNA extraction. Nucleic acid extraction was performed as described by Bostrom et al. (2004). Briefly, the filter paper was suspended in lysis buffer (400 mM NaCl, 750 mM sucrose, 20 mM ethylenediaminetetraacetic acid (EDTA), 50 mM Tris-HCl pH 9.0), followed by the addition of lysozyme (1 mg mL^{-1} , final concentration) and then incubated at 37°C for 30 minutes. Next, sodium dodecyl sulfate (1%, final concentration) and proteinase K ($100\text{ }\mu\text{g mL}^{-1}$, final concentration) were added and the mixture was incubated at 55°C overnight. Filter paper was washed with TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) in another 1.5 mL tube, and lysates from both tubes were combined together before the addition of Baker's yeast tRNA (50 μg per sample). DNA was then precipitated with 1/10 volume of 3M potassium acetate (pH 5.2) and 0.6 volume isopropanol before incubation at -20°C for an hour.

After incubation, the sample was centrifuged at 20,000 g at 4°C for 20 minutes. The supernatant was poured out, and the pellet was washed with 70% ethanol and centrifuged again, before dissolving in TE buffer and kept at -20°C . DNA was extracted twice with phenol/ chloroform/ isoamyl alcohol (25:24:1, volume) before precipitation with ethanol. Extracted nucleic acid was stored in TE buffer and kept at -20°C until use.

Polymerase chain reaction (PCR) Amplification

Partial archaeal 16S rDNA gene was amplified using the universal archaea primer combination Arch21F-958R (Llirós et al., 2010). The primer sequences are shown in Table 2.1. PCR was performed with a thermal cycler (Applied Biosystems, 2720 Thermal Cycler, Singapore) using the conditions shown in Table 2.2, for 30 cycles.

PCR reaction mixture consisted of 0.2 μ M of each primer, 0.2 mM dNTP, 1x PCR buffer (35 mM Tris-HCl, pH 9.0, 25 mM KCl, 3.5 mM MgCl₂) and 2.5 U Taq polymerase (Intron, Korea). The volume of each reaction was 15 μ L. PCR product was verified and its molecular weight estimated by agarose gel (1% w/v) electrophoresis with a 0.1-10.0 kb DNA molecular weight marker (NewEngland's Biolabs, Germany). PCR product from separate tubes was then pooled to minimize PCR bias and was purified (Qiagen, Germany) prior to cloning.

Table 2.1 Universal Archaea primer sequences used in this study.

Primer		Sequence (5'-3')
Arch21F		TTC CGG TTG ATC CYG CCG GA
Arch958R		YCC GGC GTT GAM TCC AAT T
Where;	Y = C or T	M = A or C

Table 2.2 PCR conditions used in this work. Conditions used for PCR amplification during pre-cloning and plasmid PCR are similar.

	Temperature	Duration
	(°C)	(minute)
Initial denaturation	94	1
Denaturation	94	1
Annealing	56	1
Extension	72	2
Final extension	72	10

Construction of Clone Library

Amplicon was ligated into the pGEM-T vector (Promega, USA) according to the manufacturer's instruction. Ligation products were transformed into competent *E. coli* cells (Top10, Invitrogen, Canada). Transformants were screened using blue/white selection on Luria-Bertani (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride, 1.5% agar w/v) plates containing ampicillin ($100\ \mu\text{g mL}^{-1}$), X-Gal (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) ($80\ \mu\text{g mL}^{-1}$) and IPTG (isopropyl-b-D-thiogalactopyranoside) (0.5 mM final concentration), and incubated overnight (16-24 hours) at 37°C .

Plasmid Extraction

White colonies were randomly picked and plasmid DNA was extracted according to the plasmid DNA mini-preparation method of Kotchoni et al. (2003). White colonies were picked and suspended in sterile distilled water. The bacterial cells were pelleted by centrifugation at 14,000 rpm for 30 seconds. Supernatant was discarded and the pellet was resuspended in Solution I (50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0) and incubated for 5 minutes at room temperature. Freshly prepared lysis solution (0.2M NaOH, 1% (w/v) sodium dodecyl sulphate (SDS)) was added, and mixed gently to avoid breaking the plasmid. This solution was then incubated on ice for 5 minutes. Then, 3M ice-cold potassium acetate solution was added to precipitate bacterial proteins, cell debris and SDS. The solution was allowed to precipitate for 5 minutes on ice. The mixture was then centrifuged at 14,000 rpm for 5 minutes at 4°C . Supernatant were transferred into another tube where 2 volumes of absolute ethanol were added, vortexed, and kept at room temperature for 2 minutes. The plasmid DNA was precipitated in the pellet by

centrifugation at 14,000 rpm for 5 minutes at 4°C. The pellet was then washed with ice cold 70% (v/v) ethanol, mixed well and centrifuged again. Supernatant was discarded. The pellet was allowed to air-dry to remove residual ethanol. Pellet was resuspended in 25 µL of TE buffer with 20 µg mL⁻¹ RNase and incubated at 37°C for 5 minutes to digest away all contaminating *Escherichia coli* RNA. The suspension was then kept at -20°C for further use. An aliquot of the extracted plasmid were checked in 0.8% (w/v) agarose gel.

Purified plasmids were confirmed for insert of the right size (~900 bp) using the Arch 21f/958r primers as described above. The amplicons were checked by running agarose gel (1% w/v) electrophoresis with a 0.1-10.0 kb DNA ladder (New England Biolabs, Germany).

Restriction Fragment Length Polymorphism (RFLP) Analysis

Restriction fragment length polymorphism (RFLP) analysis of 16S rDNA clones were used to define phlotypes (Dunbar et al., 1999). The PCR-amplified products of positive recombinants were digested separately using *RsaI*, *CfoI* and *DdeI* (Roche, USA) restriction enzymes for 2 hours at 37°C. Restriction fragments were resolved in agarose (2.25% w/v) gel electrophoresis and stained with ethidium bromide. Clones that produced identical patterns with three restriction enzymes were grouped into a single RFLP phylotype.

At least one clone that was representative of each RFLP phylotype was partially sequenced using the M13 forward primer (5'-TGTAACGACGGCCAGT-3'). Sequencing results were trimmed with Bioedit

Sequence Alignment Editor Version 7.0.5.3 (Hall, 1999). All sequences were checked for possible chimera formation using the online chimera-checking tool, DECIPHER (Wright et al., 2011). The nearest neighbours were retrieved from the NCBI database through the BLAST search (Altschul et al., 1990). Sequences were also compared to the Ribosomal Database Project II (Cole et al., 2007) for further confirmation of the sequence identity. The 97% cut-off for sequence similarity is used to delimit an OTU (Galand et al., 2006, Yan et al., 2006, Zhang et al., 2010). Sequences of unique OTUs were deposited in the GenBank database under the following accession numbers: JQ 415912 – JQ 415914, JQ 626832 – JQ 627017 and JX 103571 – JX 103597 (Appendix B).

Diversity and phylogenetic analysis

Library coverage (C) was calculated by the method of Good; $C = [1 - (n / N)] \times 100$, where n is the number of OTUs in a sample represented by one clone and N is the total number of clones in a library (Good, 1953). Rarefaction analysis, Shannon-Wiener diversity index (H') and two richness estimators, the abundance-based coverage estimator (S_{ACE}) and the bias-corrected Chao1 (S_{Chao1}), were calculated using the program MOTHUR version 1.23.1 (Schloss et al., 2009). The calculations for the Chao1 richness estimator are as follow:

$$S_{chao1} = S_{obs} + [n_1 (n_1 - 1) / 2 (n_2 + 1)]$$

where,

S_{chao1} = the estimated richness

S_{obs} = the observed number of species

n_1 = the number of OTUs with only one sequence (singletons)

n_2 = the number of OTUs with only two sequences (doubletons)

While Chao1 uses singletons and doubletons, S_{ACE} uses OTUs with one to ten clones each. The calculations for ACE richness estimator are implemented as described by Chao et al. (1993). Evenness (E) was calculated from the Shannon-Wiener diversity index (H') where $E = H' / \ln (S)$, where S is the total number of OTUs (Margalef, 1958).

Phylogenetic tree was constructed by the maximum likelihood method with Jukes-Cantor substitution model using the program BOSQUE version 1.7.152

(Ramírez-Flandes and Ulloa, 2008). The confidence values of the phylogenetic tree were obtained through bootstrap analysis of 1000 trial replications. The phylogenetic relatedness of archaeal 16S rDNA community was compared by weighted non-normalized principal coordinate analysis (PCoA) using the Quantitative Insights Into Microbial Ecology (QIIME) software version 1.5.0 (Caporaso et al., 2010).

Analysis of variance (ANOVA) was carried out to determine whether there is any difference in both abiotic and biotic factors. Differences among the four sites were analyzed with multivariate analysis of variance (MANOVA). Bacterial abundance was log-transformed before statistical analysis. Correlations between OTUs and environmental factors were analysed by canonical correspondence analysis (CCA). All statistical analysis was performed using the program PAST version 2.14 (Hammer et al., 2001), unless otherwise stated. A flowchart showing the experimental design followed in this project is shown in Figure 2.1.

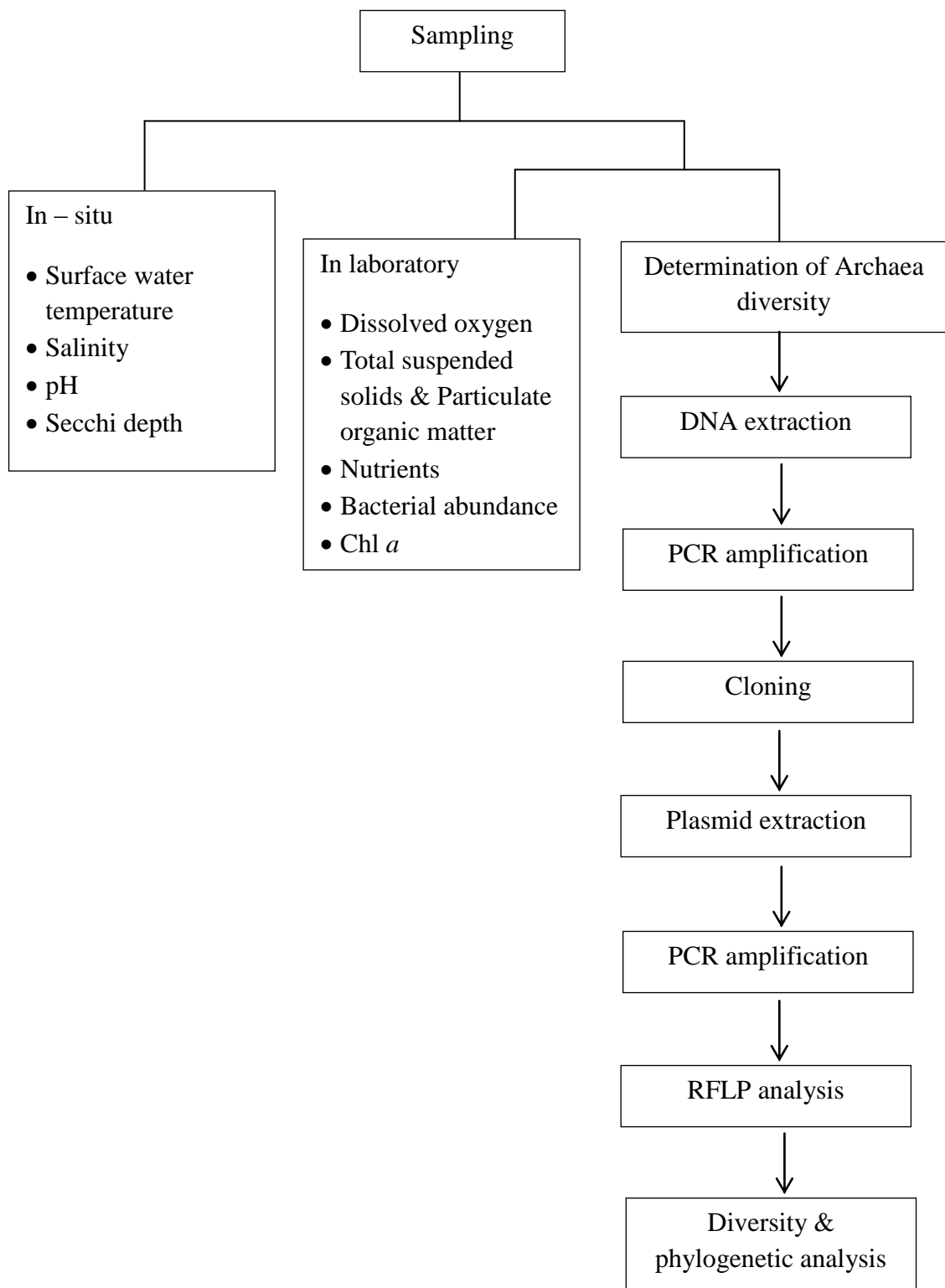


Figure 2.1 Flowchart of experimental methodology.

RESULTS

3.0 Physico-chemical characteristics

Seawater temperature was relatively stable (CV, 2.0 – 9.3%), ranging from 29.4 to 29.9 °C with an average of 29.6°C whereas salinity measured ranged from 3.7 to 27.5 ppt (Table 3.0). Salinity fluctuated over a wider range in KS and SM (CV, >90%) than in PK and PD (CV, <10%). Analysis of variance (ANOVA) showed that salinity was significantly different ($F=17.58$, $df=16$, $p < 0.001$) with Sungai Muar showing the lowest value (3.7 ± 3.6 ppt). Salinity at SM was different from that at Port Dickson (27.5 ± 1.4 ppt; $q=8.15$, $p < 0.001$) and Port Klang (26.8 ± 1.2 ppt; $q=7.90$, $p < 0.001$). Salinity levels at Kuala Selangor ranged from 4.10 to 25.50 ppt with an average of 10.3 ± 9.4 ppt which was significantly different from Port Dickson which ranged from 26.1 to 28.9 ppt (27.5 ± 1.4 ppt; $q=5.89$, $p < 0.01$) and Port Klang ranged from 25.5 to 27.9 ppt (26.8 ± 1.2 ppt; $q=5.64$, $p < 0.01$).

There was no significant difference in surface water temperature and water clarity or turbidity, as measured with Secchi disc among the four sites. Highest surface seawater pH levels was observed in Port Dickson, ranging from 7.83 to 8.02 (average, 7.91 ± 0.10) and was different ($F=5.03$, $df=14$, $p < 0.05$) when compared to Sungai Muar, ranging from 5.31 to 7.44 (6.5 ± 0.8 ; $q=4.54$, $p < 0.05$).

Highest average concentration of dissolved oxygen was observed in Port Dickson, ranging from 204 to 209 μM (206 ± 2 μM ; $F=5.95$, $df=11$, $p < 0.05$). This was significantly different from Port Klang, ranging from 135 to 161 μM (149 ± 13

μM ; $q=4.79$, $p < 0.05$) and Sungai Muar, ranging from 127 to 154 μM (141 ± 13 ; $q=5.48$, $p < 0.05$). Dissolved oxygen in Kuala Selangor ranged from 135 to 205 μM with an average of 163 ± 37 μM .

Although Kuala Selangor has the highest average total suspended solids (TSS) (110.3 ± 116.5 mgL^{-1}) and particulate organic matter (POM) (16.5 ± 13.0 mgL^{-1}), they were not significantly different from the other three sites sampled. Highest average concentration of ammonium was observed in Sungai Muar (8.25 ± 3.56 μM ; $F=4.5$, $df=11$, $p < 0.05$). This was significantly different from that at Port Dickson (0.67 ± 0.14 μM ; $q=4.79$, $p < 0.05$). High fluctuation of TSS was observed at Kuala Selangor, ranging from 42.4 mgL^{-1} to 244.8 mgL^{-1} . TSS at Sungai Muar ranged from 20.4 mgL^{-1} to 125.6 mgL^{-1} . Average TSS (Table 3.0) was 51.9 ± 14.9 mgL^{-1} , 66.4 ± 4.5 mgL^{-1} and 76.5 ± 53.0 mgL^{-1} at Port Dickson, Port Klang and Sungai Muar respectively.

Nitrite concentration was lowest at Port Dickson (0.24 ± 0.21 μM ; $F=11.39$, $df=11$, $p < 0.01$) which was significantly different from Port Klang (5.44 ± 0.97 μM ; $q=7.28$, $p < 0.01$). Nitrite concentration at Port Klang ranged from 4.41 μM to 6.65 μM with an average of 5.44 ± 0.97 μM . This was different from nitrite concentration recorded at Sungai Muar (0.75 ± 0.38 μM ; $q=6.56$, $p < 0.01$).

Kuala Selangor showed relatively higher average silicate concentration (41.75 ± 18.71 μM) compared to Port Dickson (10.69 ± 7.26 μM), Port Klang (28.33 ± 8.32 μM) and Sungai Muar (32.22 ± 8.34 μM). Although this difference was not statistically significant, Tukey's test shows a significant difference of silicate concentration in Port Dickson when compared to Kuala Selangor (32.22 ± 8.34 μM ;

$q=4.62$, $p < 0.05$). Nitrate, urea and phosphate concentrations also did not show any statistical (ANOVA) differences among the four sites ($p < 0.05$).

Although there was no significant difference for the average total bacteria abundance (Table 3.0) among the four sites, spatial variation of chlorophyll *a* concentrations was highly significant ($F=49.44$, $df=11$, $p < 0.001$) among all four sites. Highest chlorophyll *a* concentrations was measured at Kuala Selangor ($16.4 \pm 2.3 \mu\text{gL}^{-1}$), and was higher than that measured at Port Dickson ($3.1 \pm 0.4 \mu\text{gL}^{-1}$; $q=13.78$, $p < 0.001$), Port Klang ($3.2 \pm 1.0 \mu\text{gL}^{-1}$; $q=13.7$, $p < 0.001$) and Sungai Muar ($2.3 \pm 2.2 \mu\text{gL}^{-1}$; $q=14.64$, $p < 0.001$).

Individual environmental variables separated the four sampling sites via Tukey's test with the exception of silicate ($P > 0.05$) and total bacterial abundance ($P > 0.05$). Based on the same environmental variables, MANOVA showed the four sites were significantly different (Wilk's Lambda $F = 36.32$, $P = < 0.01$).

Table 3.0 Environmental parameters at KS, PK, PD and SM. Values are means (\pm standard deviations). *, **, *** showed the results of ANOVA testing at $P < 0.05$; $P < 0.01$ and $P < 0.001$, respectively. The same letters of the alphabet were used to indicate values whose means were significantly different.

Parameters	Station			
	KS	PK	PD	SM
Salinity (ppt)***	10.3 \pm 9.4 ^{ab}	26.8 \pm 1.2 ^{bd}	27.5 \pm 1.4 ^{ac}	3.7 \pm 3.6 ^{cd}
pH*	7.0 \pm 0.3	7.6 \pm 0.1	7.9 \pm 0.1 ^a	6.5 \pm 0.8 ^a
DO (μ M)*	163 \pm 37	149 \pm 13 ^a	206 \pm 2 ^{ab}	141 \pm 13 ^b
Temperature ($^{\circ}$ C)	29.9 \pm 1.3	29.4 \pm 0.6	29.6 \pm 0.7	29.6 \pm 2.7
TSS (mgL ⁻¹)	110.3 \pm 116.5	66.4 \pm 4.5	51.9 \pm 14.9	76.5 \pm 53.0
POM (mgL ⁻¹)	16.5 \pm 13.0	9.9 \pm 0.1	6.8 \pm 3.6	13.9 \pm 7.7
Secchi depth (m)	0.3 \pm 0.1	0.6 \pm 0.1	1.1 \pm 0.7	0.3 \pm 0.1
Ammonium(μ M)*	7.12 \pm 4.12	6.09 \pm 0.65	0.67 \pm 0.14 ^a	8.25 \pm 3.56 ^a
Nitrate (μ M)	13.37 \pm 10.56	6.37 \pm 0.53	1.77 \pm 1.47	15.83 \pm 4.10
Nitrite(μ M)**	3.30 \pm 2.16	5.44 \pm 0.97 ^{ab}	0.24 \pm 0.21 ^a	0.75 \pm 0.38 ^b
Urea (μ M)	1.19 \pm 0.40	0.58 \pm 0.24	0.31 \pm 0.33	1.13 \pm 0.59
Phosphate (μ M)	0.52 \pm 0.16	1.26 \pm 0.01	1.44 \pm 1.03	0.39 \pm 0.32
Silicate (μ M)	41.75 \pm 18.71 ^a	28.33 \pm 8.32	10.69 \pm 7.26 ^a	32.22 \pm 8.34
Bacterial abundance (10 ⁶ cells ml ⁻¹)	2.1 \pm 1.1	2.4 \pm 0.5	1.3 \pm 0.2	1.5 \pm 0.4
Chl <i>a</i> (μ gL ⁻¹)***	16.4 \pm 2.3 ^{abc}	3.2 \pm 1.0 ^a	3.1 \pm 0.4 ^b	2.3 \pm 2.2 ^c

3.1 Molecular results

Archaeal DNA was successfully extracted and amplified from the seawater samples, and an archaeal 16S rDNA library was constructed for each sampling site. A total of 1265 randomly chosen clones were picked and its plasmid extracted (Figure 3.0). Of that total, only 1181 clones contained insert of the right size (Figure 3.1). After RFLP analysis with *RsaI* (Figure 3.2), *CfoI* (Figure 3.3) and *DdeI* (Figure 3.4) restriction enzymes, 76, 48, 45 and 91 distinct RFLP phylotypes were identified for KS, PK, PD and SM clone libraries, respectively.

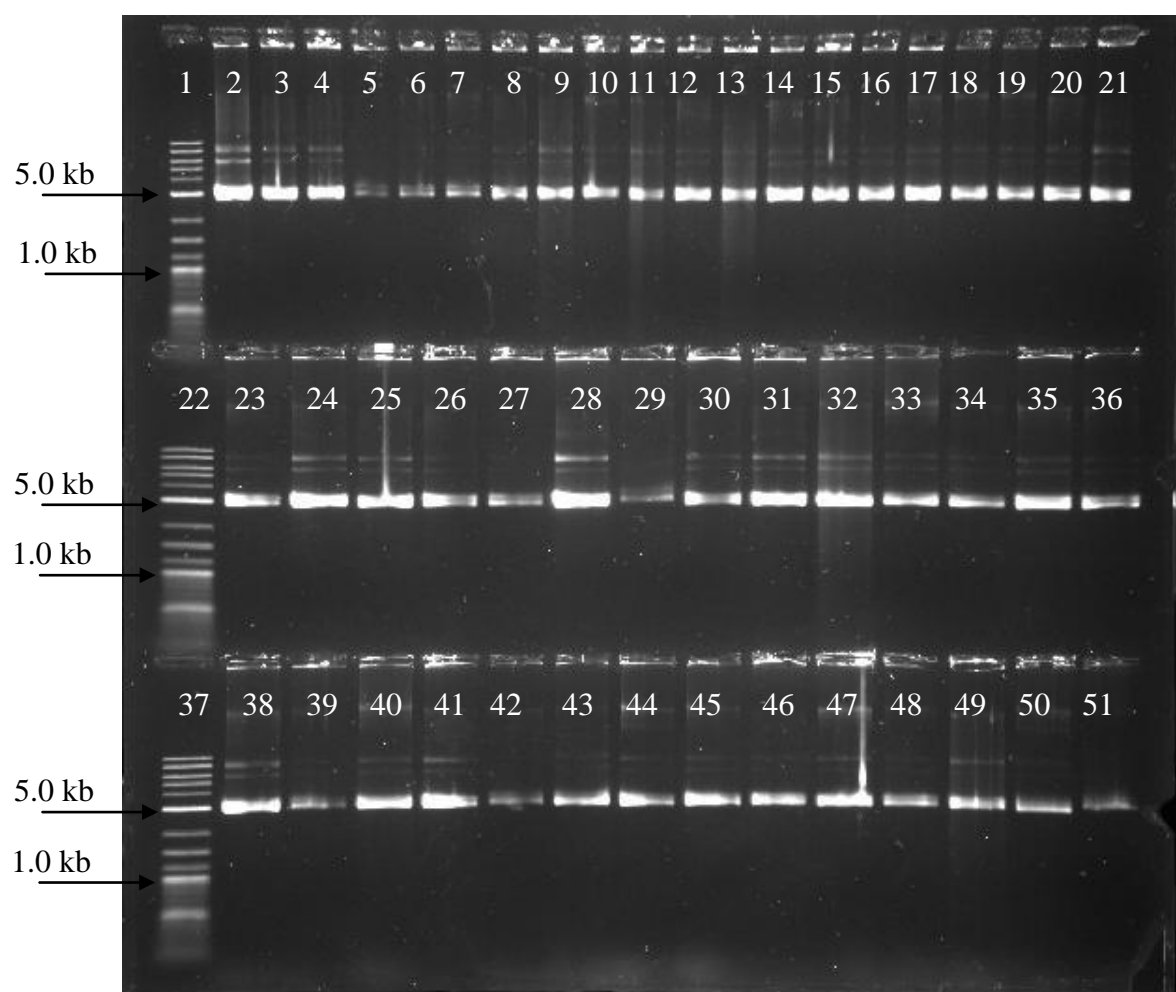


Figure 3.0 Gel electrophoresis of extracted plasmids. Lanes 1, 22, 37: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs). Lanes 2-21, 23-36, 38-51: extracted plasmids.

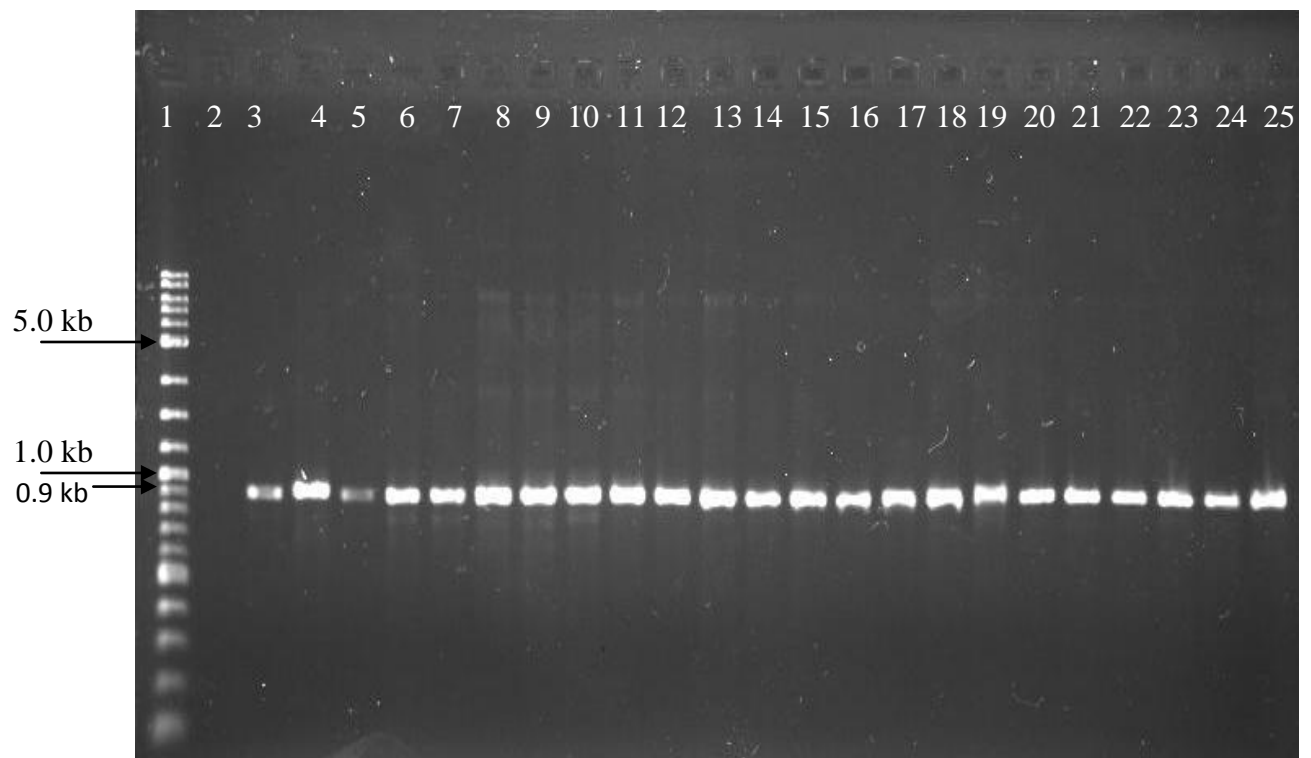


Figure 3.1 Gel electrophoresis of PCR amplifications of DNA fragments from purified plasmid DNA using Arc 21f/958r primers. Lane 1: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs), Lane 2: Negative control without DNA template, Lane 3-25: Amplicons (~ 900 bp).

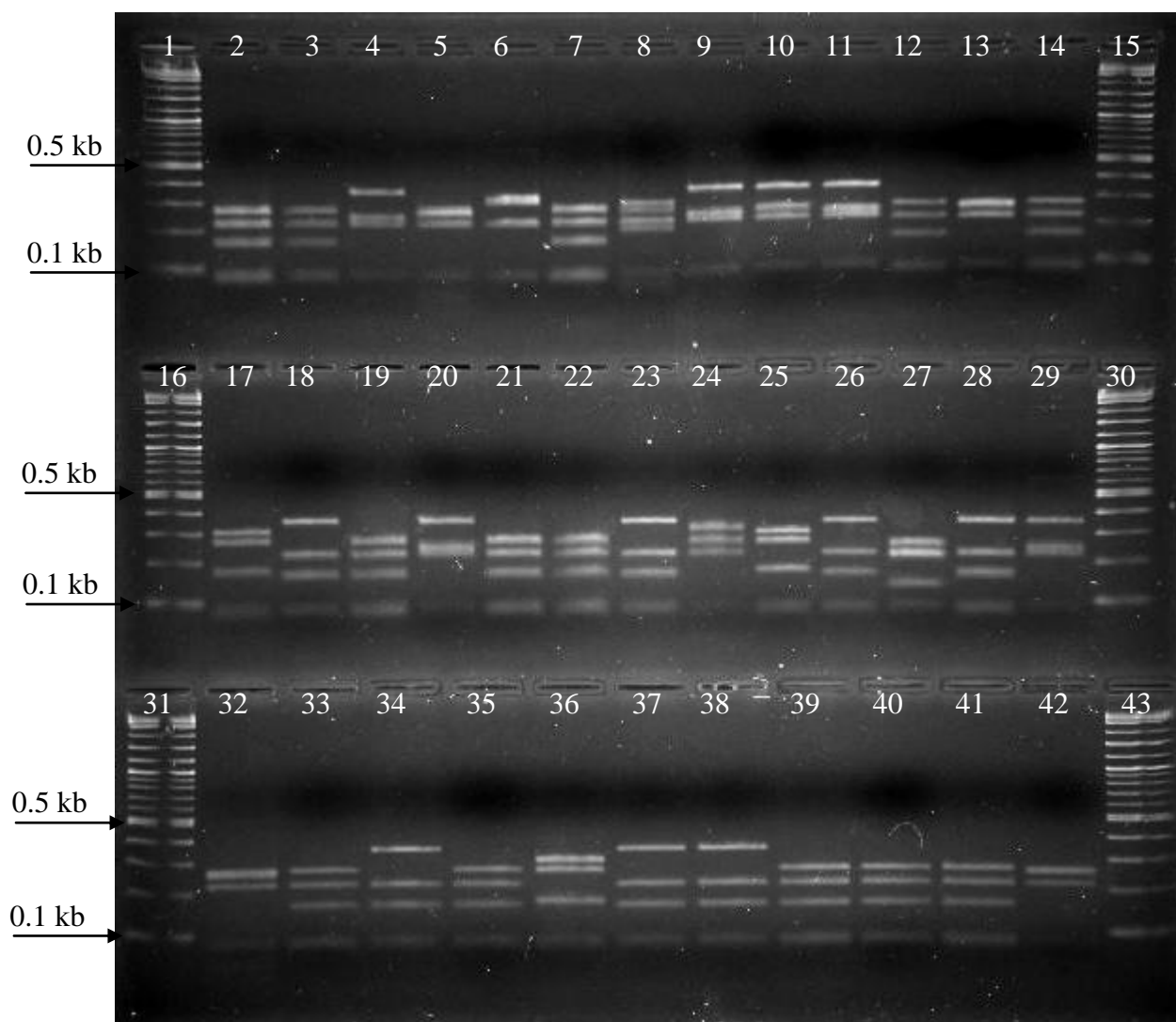


Figure 3.2 Comparison of *RsaI* RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

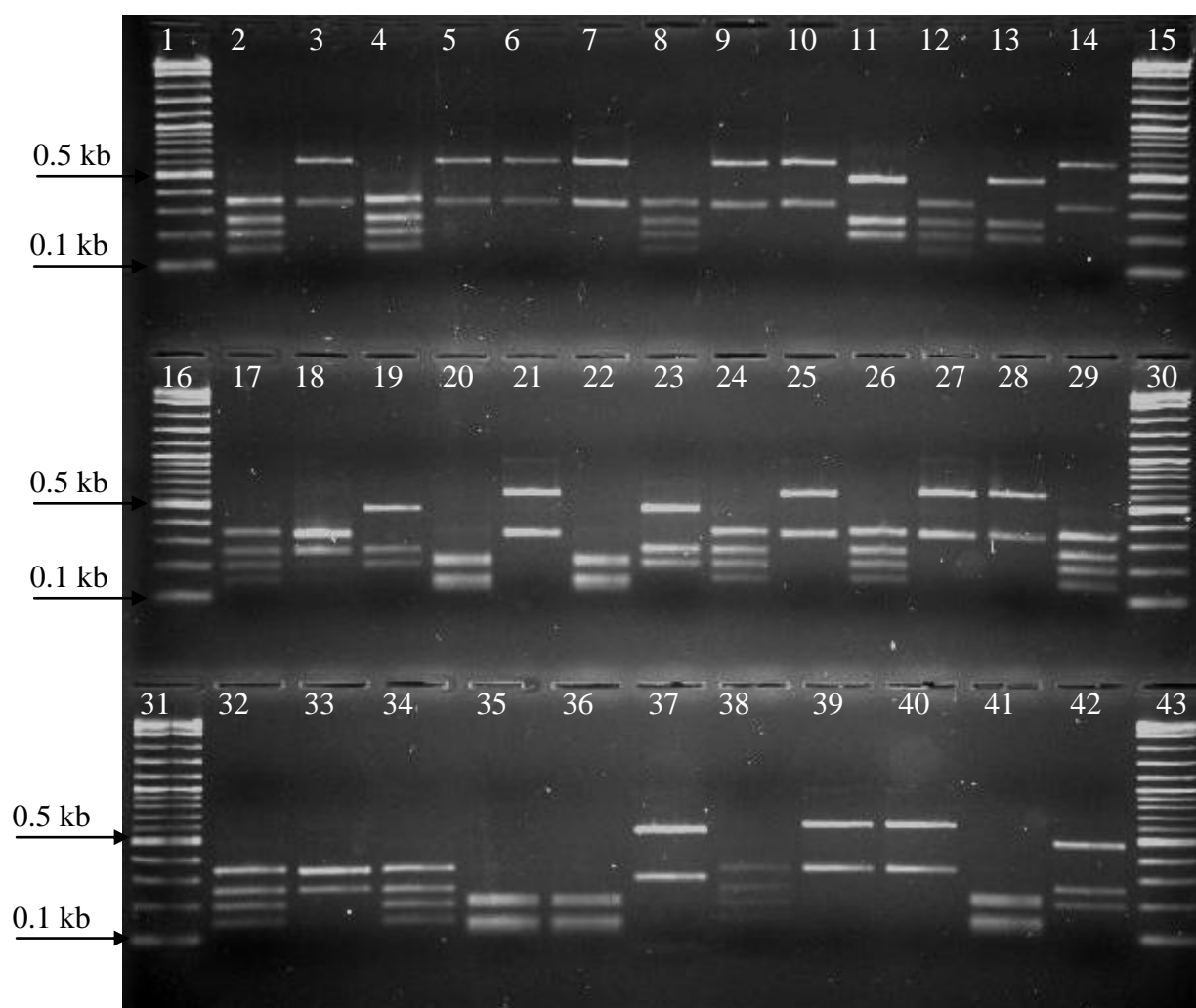


Figure 3.3 Comparison of *Cfo*I RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

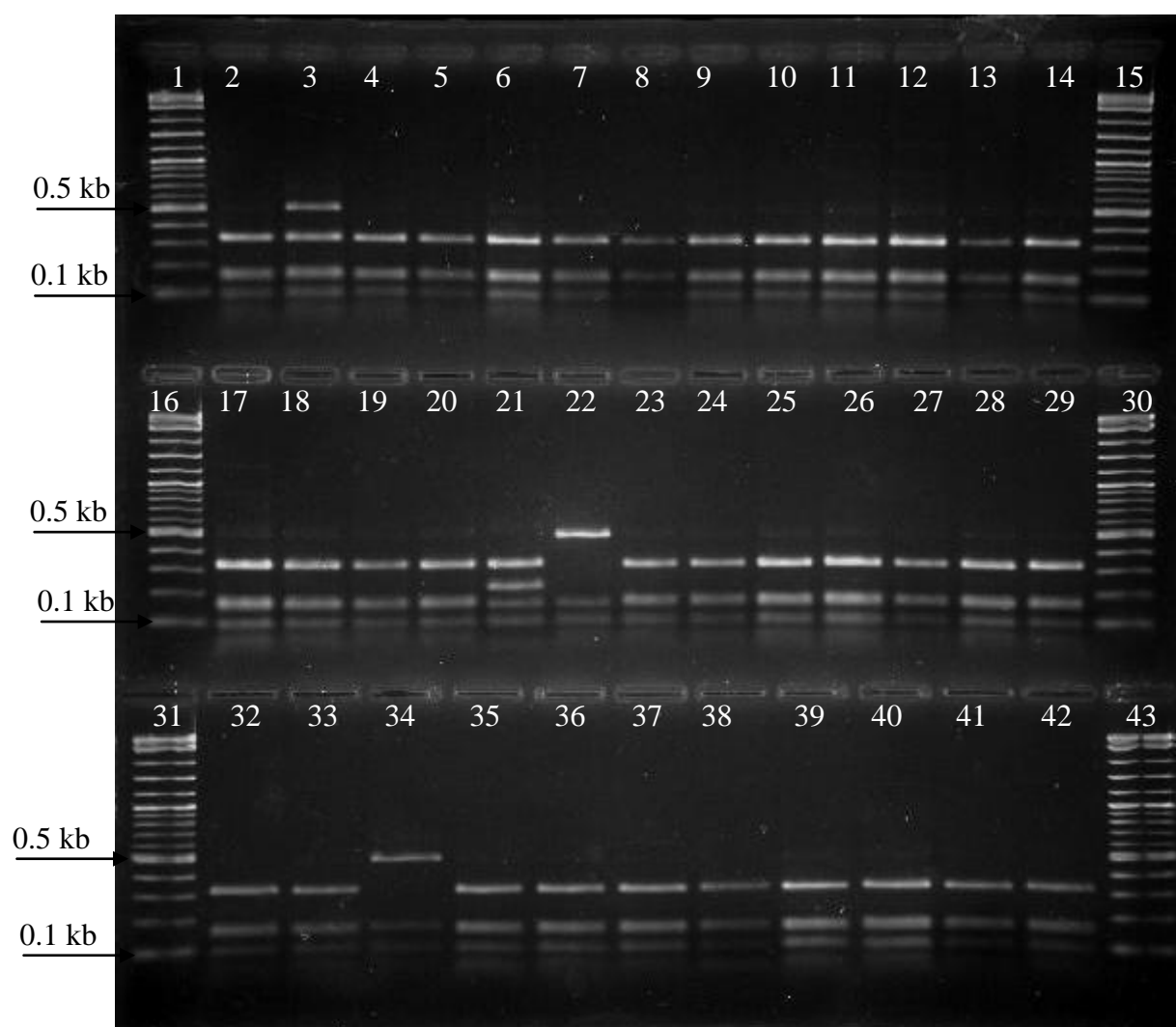


Figure 3.4 Comparison of *DdeI* RFLPs of the insert obtained from clones via PCR (lanes 2-14, 16-29, 32-42). Migration was performed on a 2.25% agarose gel. Lanes 1, 15, 16, 30, 31, 43: 0.1-10.0 kb DNA molecular weight marker (New England Biolabs).

3.2 Archaea diversity and phylogenetic relationship

Representative clone for each RFLP genotype was sequenced and chimeric artifacts were recognized and removed from further analyses. All sequences analysed were related to archaeal sequences. Unique RFLP phylotypes from each clone library were aligned using BOSQUE and 97% cut-off for sequence similarity among each RFLP phylotype was used to delimit an OTU. A total of 129 OTUs that represented 1074 clones were used in diversity analyses.

Rarefaction curve indicated that PD and PK archaeal clone libraries were representative of the archaeal communities as the rarefaction curves were approaching plateaus (Figure 3.5). Although the rarefaction curve of SM and KS were still on an incline, archaeal diversity was clearly different. By means of Shannon-Weiner diversity index, SM clone library had the highest diversity of the archaeal OTUs whereas PK clone library had the lowest diversity (Table 3.1). The number of OTUs, Chao1 and Ace richness estimates also indicated that SM had a greater Archaea diversity compared to KS, PK and PD. Although the archaeal diversity in KS and SM probably required more exhaustive sampling, Good's coverage values elucidated that more than 87% of archaeal sequence types were obtained in all clone libraries (Table 3.1), and showed that the libraries in this study had captured the majority of the archaeal OTUs. Majority of the archaeal OTUs belong to rare species represented by only a few or a single clone. Singletons i.e. OTU sequences that occur only once in each library represented 54.3%, 41.4%, 37.5% and 56.5% of the OTUs in KS, PK, PD and SM libraries, respectively. PCoA analyses showed that the archaeal communities of the four sites were clearly separated from each other. The first two principal coordinates (P1 and P2) explained 98.0 % of the total community variability.

Clone library analysis

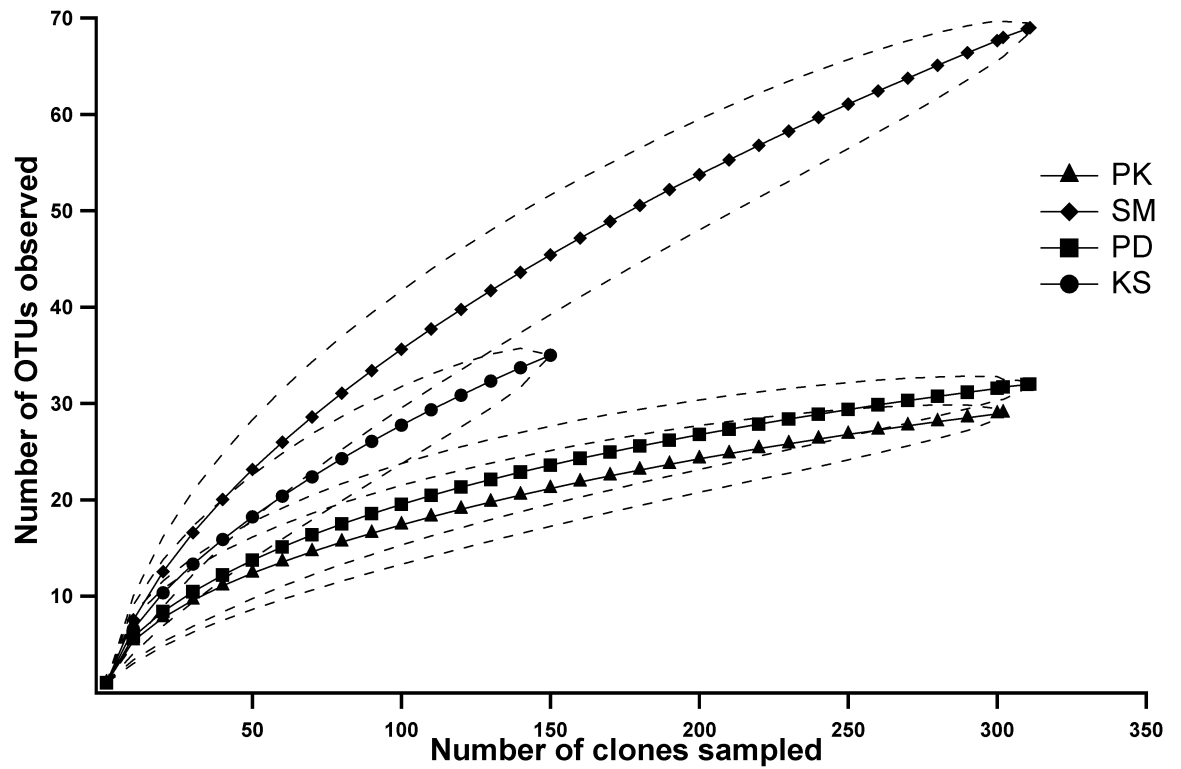


Figure 3.5 Rarefaction curve of archaeal clone libraries constructed with MOTHUR using 10,000 iterations. Dotted lines shows the higher and lower confidence intervals of each curve.

Table 3.1 Coverage, observed OTUs, richness and diversity indices for clone libraries constructed. Values in parentheses represent lower and higher confidence interval.

Clone libraries	No. of valid clones analysed	Coverage good (%)	OTUs	Richness index		Diversity index	
				Chao1 (95% CI)	Ace (95% CI)	Shannon-Wiener, H' (95% CI)	Evenness E
Kuala							
Selangor	150	87	35	54 (41-92)	118 (84-175)	2.65 (2.43-2.86)	0.75
Port Klang	302	96	29	40 (31-70)	45 (34-82)	2.16 (2.01-2.31)	0.64
Port Dickson	311	96	32	40 (34-63)	45 (36-72)	2.33 (2.18-2.48)	0.67
Sungai Muar	311	88	69	123 (92-193)	223 (173-297)	3.24 (3.07-3.40)	0.77

Relative to archaeal 16S rDNA sequences available in GenBank, sequence similarity ranged from 83% to 100%. Our sequences were closely matched with sequences retrieved from a variety of terrestrial, freshwater, estuarine, coastal and marine environments. One OTU (73 clones) was related to the sequence isolated from marine sponge and 3 OTUs, were similar to sequences isolated from corals such as *Siderastrea stellata* (2 clones) and *Alcyonium gracillimum* (173 clones). Most of the OTU sequences were related to uncultured archaeal clones than with isolated pure cultures (Table 3.2), reflecting the large number of archaeal species that remains to be cultivated (Vieira et al., 2007). Only 12 OTU sequences (69 clones) were similar ($\geq 95\%$) to nine cultured species; *Methanococcoides methylutens* (98%), *Methanosaeta concilii* (99%), *Methanoplanus paludicola* (95%), *Methanoplanus petrolearius* (97%), *Methanoplanus sp.* (97%), *Methanobacterium beijingense* (98%), *Methanolacinia paynteri* (98%), *Methanocorpusculum labreanum* (99%), *Nitrosopumilus maritimus* SCM1 (96%) and *Candidatus Nitrosopumilus sp.* NM25 (98%).

In this study, the archaeal clone sequences were affiliated with Euryarchaeota, Crenarchaeota and Thaumarchaeota, whereas no sequence was affiliated with Korarchaeota and Nanoarchaeota. Phylogenetic analysis placed the euryarchaeotal sequences into six subgroups: Methanomicrobia, Methanobacteria, Marine Group III, Marine Benthic Group – D (MBGD), Marine Group II and unaffiliated Euryarchaeota (Figure 3.7). Sequences affiliated with Euryarchaeota predominated in PD (81.3% OTUs, 75.5% clones), SM (52.2% OTUs, 50.5% clones) and PK (51.7% OTUs, 36.4% clones), respectively. The most predominant OTU from PD (82 related clones) and PK (71 related clones) was OM2, affiliated with Euryarchaeota, and exhibited 99% similarity to clone CWP-B5 obtained from the

western Pacific warm pool. Sequences similar to (95-99%) cultured methanogenic Euryarchaeota were represented by 8.7%, 5.7% and 3.4% of the OTUs in SM, KS and PK, respectively. No sequences similar to culturable Euryarchaeota were obtained from PD.

Majority of the clones (> 60.0%) in KS and PK were affiliated with Thaumarchaeota (Figure 3.6). OTUs belonging to Thaumarchaeota represented 73.3%, 61.6%, 28.6% and 24.2% of the total clones in KS, PK, SM and PD respectively. Four OTUs (60 clones) from total OTUs in all sites were affiliated with *Nitrosopumilus* sp. Of 302 archaeal 16S rDNA clones from PK, sequence OM17 was the most frequently detected phylotype (101 related clones) (Table 3.2), and this sequence is closely related (99%) to a coral-associated microbe (JF925087), and also related to sequences retrieved from the Pacific ocean (JQ226183) and the coastal waters of the Gulf of Mexico (GQ906614). These clones are of marine origin (Wright et al., 2012) and are also widely distributed in the seawater here, as shown by this study.

OTUs belonging to the Miscellaneous Crenarchaeotic Group (MCG) (Inagaki et al., 2003) represented 20.9%, 18%, 2% and 0.3% of the clones in SM, KS, PK and PD respectively. Most (78%) of the MCG OTUs were similar (90-99%) to sequences retrieved from lake sediments, estuarine sediments, tropical marine sediments and brackish sediments. Phylogenetic analysis showed the distinct clustering of Marine Group 1 (Thaumarchaeota) sequences and Miscellaneous Crenarchaeotic Group (MCG). One OTU from SM were similar (95%) to a sequence belonging to South Africa Gold Mine Crenarchaeotic Group-1 (SAGMCG-1) (Takai et al., 2001). These sequences clustered together in the phylogenetic tree (Figure 3.8).

Although the OTUs were similar (83 – 100%) to sequences from the GenBank database, some formed an outgroup in the phylogenetic tree. A separate tree was constructed for these sequences (Figure 3.9). OTUs in Clade I and Clade III were affiliated with Euryarchaeota sequences. Clade II were affiliated with MCG and OTUs in Clade IV were similar to Thaumarchaeota sequences.

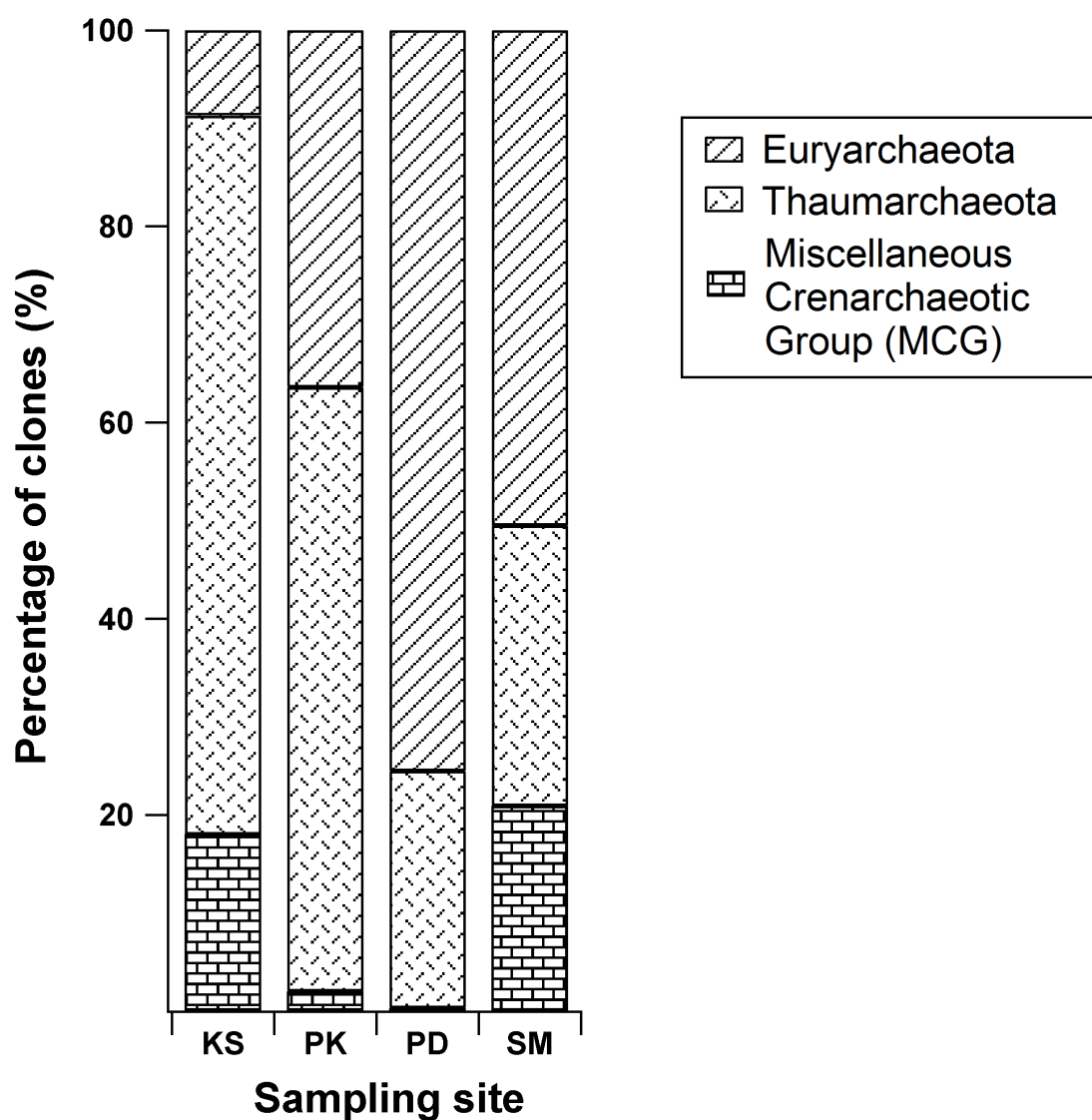


Figure 3.6 Archaeal phylotype compositions for 16S rDNA clone libraries constructed from KS, PK, PD and SM.

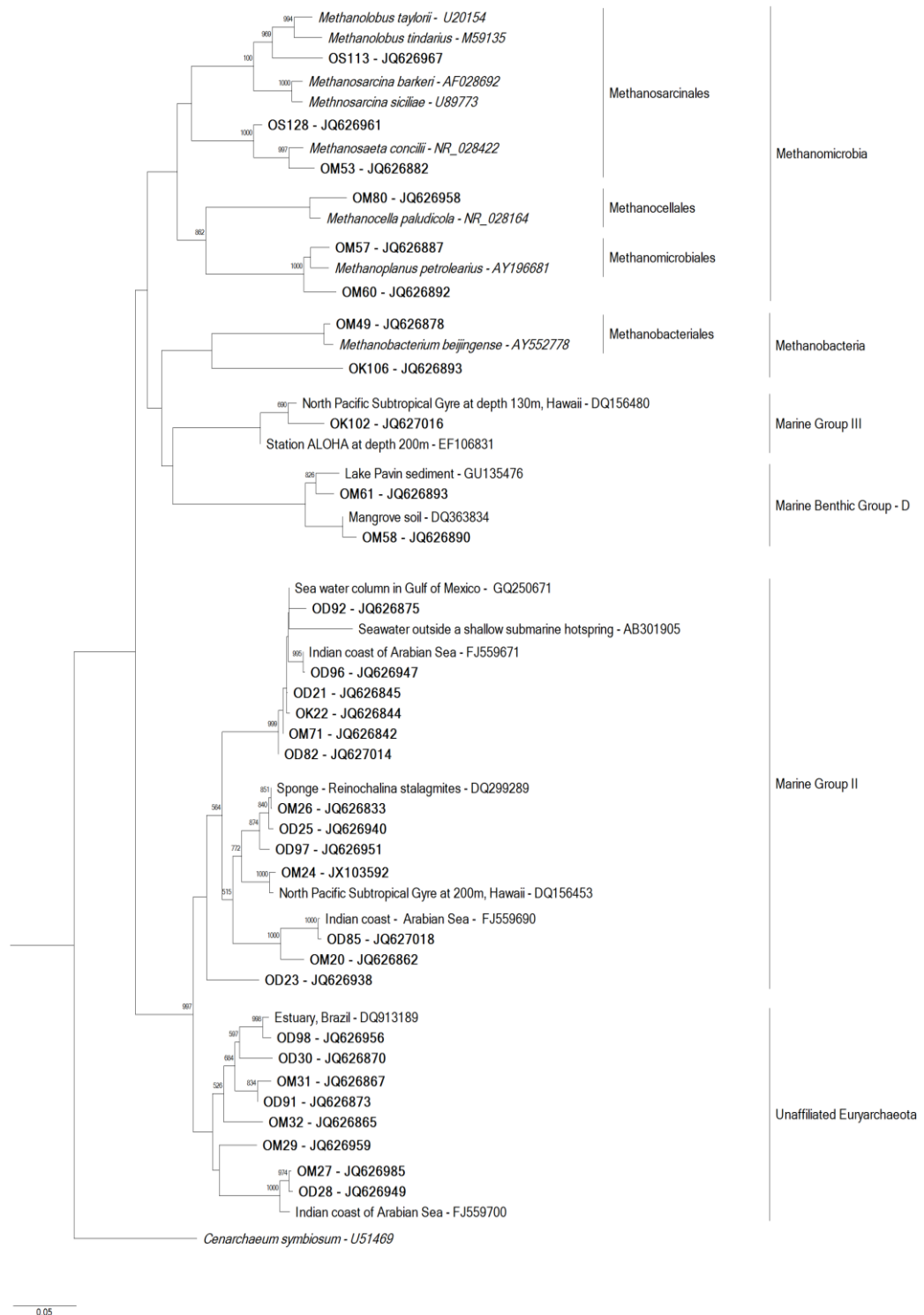


Figure 3.7 Phylogenetic analysis of Euryarchaeota 16S rDNA sequences obtained from KS, PK, PD and SM. *Cenarchaeum symbiosum* was used as an outgroup. OTUs obtained from this study were bold, and the number after the - shows the accession code deposited in GenBank. Clone sequences obtained from various environments similar to OTU sequences in this study were shown by stating the isolation source and its accession code. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

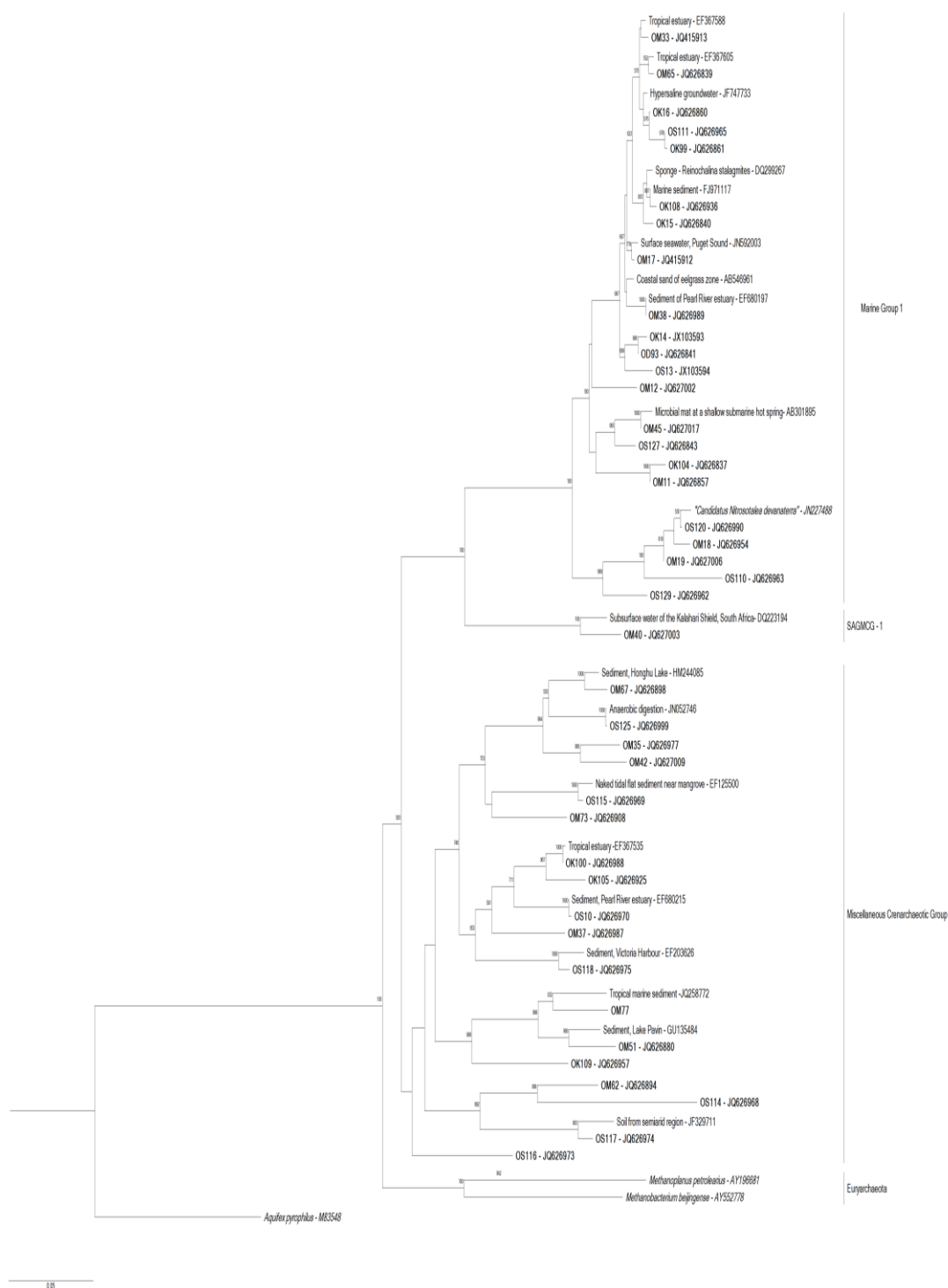


Figure 3.8 Phylogenetic analysis of Miscellaneous Crenarchaeotic Group and Thaumarchaeota (SAGMCG-1 and Marine Group 1) 16S rDNA sequences obtained from KS, PK, PD and SM. *Aquifex pyrophilus* was used as an outgroup. OTUs obtained from this study were bold, and the number after the ` - ` shows the accession code deposited in GenBank. Clone sequences obtained from various environments similar to OTU sequences in this study where shown by stating the isolation source and its accession code. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

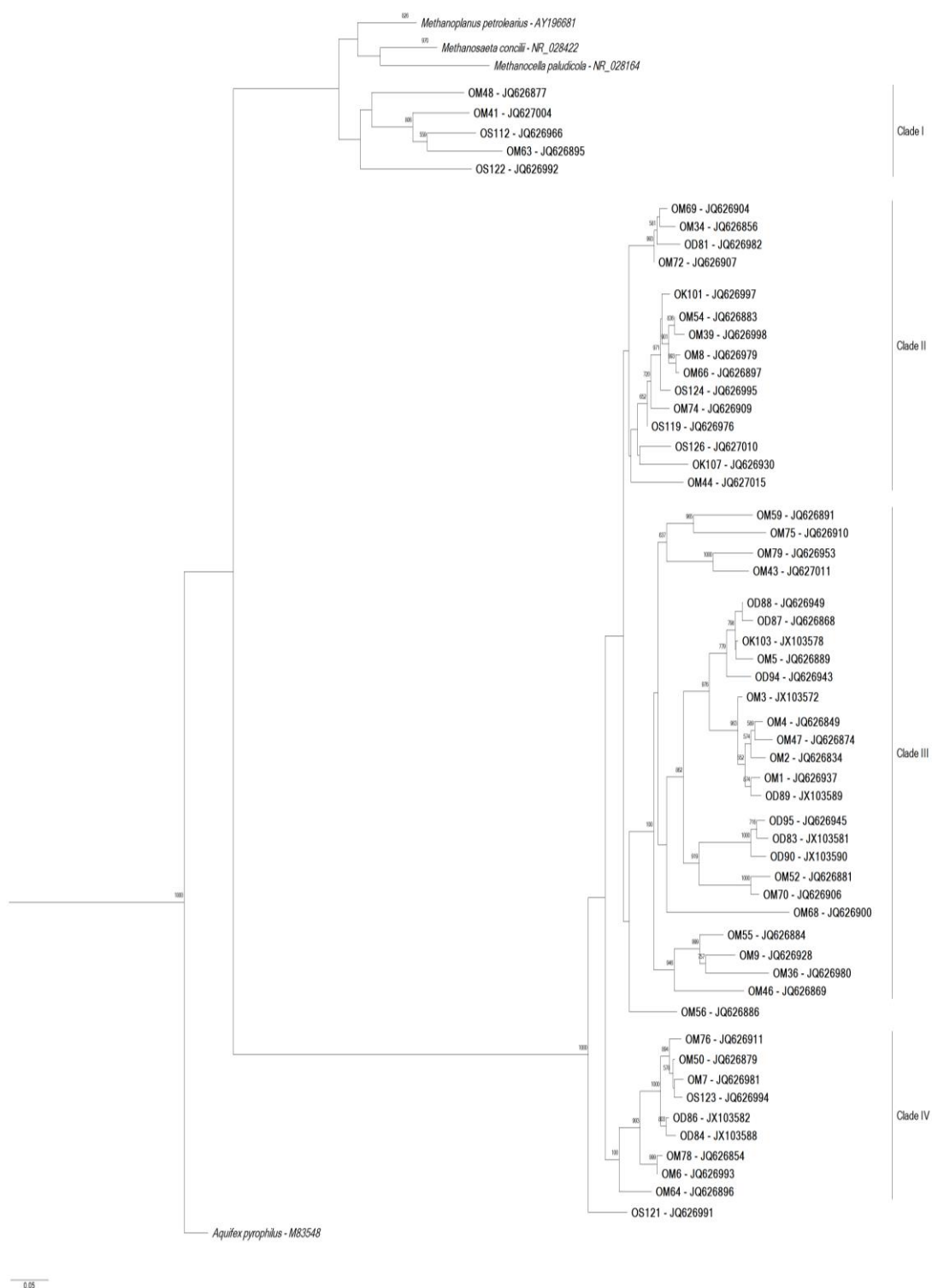


Figure 3.9 Phylogenetic analysis of archaeal 16S rDNA sequences obtained from KS, PK, PD and SM. *Aquifex pyrophilus* was used as an outgroup. OTUs obtained from this study were bold, and the number after the ` - ` shows the accession code deposited in GenBank. Sequences obtained from various environments similar to OTU sequences in this study were shown by stating the isolation source and its accession code. Italicized words represent isolated cultures. Bootstrap values of higher than 50% are shown. Scale bar represents the 5% nucleotide substitution percentage.

Table 3.2 Phylogenetic affiliations of archaeal clone libraries.

Affiliation	Clone	Abundance		Nearest neighbor	
		(KS/PK/PD/SM)	Sequence	Identity(%)	Isolated environment / Species
Euryarchaeota	OS113	1 / 0 / 0 / 0	FR733669	98	<i>Methanococcoides methylutens</i>
	OS128	1 / 0 / 0 / 0	CP002565	99	<i>Methanosaeta concilii</i>
	OM80	0 / 0 / 0 / 1	NR_028164	95	<i>Methanocella paludicola</i>
	OM57	0 / 0 / 0 / 1	CP002117	97	<i>Methanoplanus petrolearius</i>
	OM60	0 / 0 / 0 / 1	AB370246	97	<i>Methanoplanus sp</i>
	OM49	0 / 0 / 0 / 1	AY552778	98	<i>Methanobacterium beijingense</i>
	OM9	0 / 1 / 0 / 1	JQ346754	98	<i>Methanolacinia paynteri</i>
	OM36	0 / 0 / 0 / 1	CP000559	99	<i>Methanocorpusculum labreanum</i>
	OK102	0 / 2 / 0 / 0	JF715354	97	Pacific Ocean: eastern tropical south Pacific
	OD91	0 / 0 / 4 / 0	JF715299	99	“
	OM5	1 / 0 / 0 / 2	JF715343	99	“
	OD94	0 / 0 / 1 / 0	JF715377	99	“
	OD25	0 / 0 / 2 / 0	HQ529816	99	Pacific Ocean: western Pacific warm pool
	OM2	1 / 3 / 82 / 71	HQ529816	99	“

OM24	1 / 1 / 1 / 3	HQ529784	99	“
OM3	0 / 0 / 1 / 1	HQ529789	98	“
OD92	0 / 0 / 1 / 0	GQ250671	97	Sea water column in Gulf of Mexico
OD21	0 / 1 / 28 / 0	GQ250671	99	“
OK22	2 / 11 / 0 / 0	GQ250671	99	“
OM71	0 / 0 / 0 / 10	GQ250671	94	“
OD82	0 / 0 / 2 / 0	GQ250671	96	“
OM27	0 / 1 / 0 / 1	FJ559700	98	Water from the Indian coast of the Arabian Sea
OD28	0 / 2 / 6 / 0	FJ559700	97	“
OD88	0 / 0 / 1 / 0	FJ559700	97	“
OD96	0 / 0 / 8 / 0	FJ559671	98	“
OD85	0 / 0 / 3 / 0	FJ559690	99	“
OM20	0 / 0 / 8 / 2	FJ559638	99	“
OD89	0 / 0 / 1 / 0	FJ559548	99	“
OD95	0 / 0 / 6 / 0	DQ156479	99	North Pacific Subtropical Gyre (Hawaii)
OD90	0 / 0 / 1 / 0	DQ156479	98	“
OD83	0 / 0 / 1 / 0	DQ156480	100	“
OD23	0 / 0 / 2 / 0	DQ156348	99	“
OM31	4 / 8 / 0 / 4	JN591984	98	Surface seawater, Puget Sound
OM32	0 / 0 / 1 / 10	JN591984	95	“
OD30	0 / 4 / 6 / 0	AB301901	95	Seawater outside a shallow submarine hot

OM29	0 / 1 / 0 / 1	AB538535	98	Seawater at 10m depth, Japan
OM4	0 / 4 / 56 / 17	JQ227360	99	Ocean water marine sample
OM47	0 / 0 / 0 / 1	AB301905	99	Seawater outside a shallow submarine hot spring
OK103	0 / 1 / 0 / 0	DQ186704	97	Guanabara bay water (tropical)
OM1	0 / 0 / 2 / 1	AB629569	99	Bottom seawater, Japan
OD98	0 / 0 / 2 / 0	DQ913189	98	Tropical estuarine
OM48	0 / 0 / 0 / 1	FN691704	98	Lake Llebrete, Spain
OM68	0 / 0 / 0 / 2	FN691682	96	“
OM43	0 / 0 / 0 / 1	FN691738	99	“
OD87	0 / 0 / 8 / 0	FJ350099	95	Lake Pontchartrain
OK106	0 / 1 / 0 / 0	GU135476	94	Lake Pavin sediment
OM61	0 / 0 / 0 / 1	GU135476	94	“
OM41	0 / 0 / 0 / 3	HM244280	94	Lower sediment from Honghu Lake
OM75	0 / 0 / 0 / 2	HM244280	93	“
OM70	0 / 0 / 0 / 2	HM244287	99	
OM55	0 / 0 / 0 / 1	JQ245868	99	Estuarine sediment, China
OM53	0 / 0 / 0 / 1	EU244173	96	River Leine sediment
OM46	0 / 0 / 0 / 3	JF304117	99	Outfall sediment, Xiangjiang River
OM63	0 / 0 / 0 / 1	AY396650	97	Tidal flat sediment,

	OM52	0 / 0 / 0 / 1	GU127545	97	Sediment; anoxic zone
	OM58	0 / 0 / 0 / 1	DQ363834	98	Mangrove soil
	OM79	0 / 0 / 0 / 1	AB652184	100	Rice paddy soil, Japan
	OM59	0 / 0 / 0 / 1	HQ614273	96	Acidic peatland, 50cm depth
	OS112	1 / 0 / 0 / 0	AB600445	94	Iron-rich microbial mat, Japan
	OS122	1 / 0 / 0 / 0	AB600446	83	‘‘
	OM26	0 / 69 / 0 / 4	DQ299289	99	Sponge
	OD97	0 / 0 / 1 / 0	JQ611073	92	Venting fluid in white vent off Kueishan Island
Marine Grp1	OM65	0 / 0 / 0 / 14	CP000866	96	Nitrosopumilus maritimus SCM1
	OK108	0 / 3 / 0 / 0	CP000866	98	‘‘
	OK15	21 / 20 / 0 / 0	CP000866	98	‘‘
	OM38	0 / 0 / 0 / 2	AB546961	98	Candidatus Nitrosopumilus sp. NM25
	OK14	42 / 2 / 0 / 0	EF367625	99	Tropical estuary
	OM33	0 / 0 / 0 / 33	EF367588	99	‘‘
	OK104	0 / 15 / 0 / 0	FJ559501	99	Water from the Indian coast of the Arabian Sea
	OK16	21 / 3 / 0 / 0	JF747733	99	Manantial del Toro hypersaline groundwater

OM11	1 / 1 / 2 / 21	JQ245932	98	Estuarine sediment, Jiulong River
OM50	0 / 0 / 0 / 1	JQ245933	99	“
OM7	0 / 0 / 0 / 2	JQ245933	97	“
OS123	1 / 0 / 0 / 0	JQ258705	99	Tropical marine sediment
OD93	0 / 0 / 5 / 0	JQ258214	94	“
OD86	0 / 0 / 1 / 0	HQ611184	99	Marine sediments, South China Sea
OS127	2 / 0 / 0 / 0	HM171771	98	Sediment, Hainan island
OM12	2 / 2 / 0 / 2	DQ641722	99	Madovi Estuary sediment
OS129	1 / 0 / 0 / 0	DQ641790	94	“
OS13	8 / 0 / 0 / 0	EF051158	99	Mangrove sediment
OK99	0 / 39 / 0 / 0	GQ848378	90	Deep sea hydrothermal
OS110	2 / 0 / 0 / 0	GU366919	97	Temperate forest soil
OM19	2 / 0 / 0 / 1	FJ936677	100	Volcano mud
OM45	0 / 0 / 0 / 1	AB301895	99	Microbial mat at a shallow submarine hot spring, Japan
OD84	0 / 0 / 2 / 0	AB301889	99	Microbial mat at a
OS120	1 / 0 / 0 / 0	AM910765	96	Acidic fen, Germany
OM18	4 / 0 / 0 / 5	AM910765	98	“
OS111	2 / 0 / 0 / 0	HM223523	90	Coral
OM17	0 / 101 / 65 / 6	JF925087	99	“
OM76	0 / 0 / 0 / 1	JF925088	99	“

SAGMCG-1	OM40	0 / 0 / 0 / 2	DQ223194	95	Subsurface water,
MCG	OS119	1 / 0 / 0 / 0	EF367617	93	Kalahari Shield
	OK100	0 / 1 / 0 / 0	EF367535	99	Tropical estuary
	OM67	0 / 0 / 0 / 1	HM244085	97	“
	OM56	0 / 0 / 0 / 1	HM244338	96	Sediment from Honghu Lake
	OM73	0 / 0 / 0 / 1	HM244292	93	“
	OM8	0 / 0 / 0 / 8	HM244111	98	“
	OM74	0 / 0 / 0 / 1	HM244170	99	“
	OM77	0 / 0 / 0 / 5	JQ245926	94	Estuarine sediment,
	OM69	0 / 0 / 0 / 2	JQ245926	98	Jiulong River
	OS114	1 / 0 / 0 / 0	JQ245936	98	“
	OM39	0 / 0 / 0 / 1	JQ245908	98	“
	OM78	0 / 0 / 0 / 22	JQ245932	97	“
	OM6	10 / 0 / 0 / 2	JQ245932	99	“
	OK101	0 / 2 / 0 / 0	JQ245917	99	“
	OS126	5 / 0 / 0 / 0	JQ245879	99	“
	OS116	1 / 0 / 0 / 0	JQ245901	96	“
	OM34	0 / 0 / 0 / 4	JQ245942	97	“
	OS124	1 / 0 / 0 / 0	JQ258733	99	Tropical marine sediment
	OK107	0 / 1 / 0 / 0	JQ258647	99	“
	OM37	0 / 0 / 0 / 1	JQ257967	98	“
	OK105	0 / 1 / 0 / 0	FR852573	98	Brackish marine
	OK109	0 / 1 / 0 / 0	GQ927545	90	Marine sediment from the tropical western

OM51	0 / 0 / 0 / 1	GU135484	94	Lake Pavin sediment
OM44	0 / 0 / 0 / 1	GU135493	98	“
OS10	2 / 0 / 0 / 0	EF680215	99	Sediment, Pearl River Estuary
OS118	1 / 0 / 0 / 0	EF203626	97	Sediments of Pang Chau, west side of Victoria
OM62	0 / 0 / 0 / 1	HM171870	97	Sediment, Hainan island
OM35	0 / 0 / 0 / 6	EU247265	96	Lake Coeur d'Alene
OM72	0 / 0 / 0 / 2	GU127570	93	Sediment; anoxic zone
OM42	0 / 0 / 0 / 1	JF304137	98	Outfall sediment, Xiangjiang River
OS115	2 / 0 / 0 / 0	EF125500	98	Naked tidal flat sediment near mangrove
OD81	0 / 0 / 1 / 0	EF125521	99	Mangrove soil
OS117	1 / 0 / 0 / 0	JF329711	97	Soil from semiarid region
OM54	0 / 0 / 0 / 1	JF329842	99	“
OM64	0 / 0 / 0 / 1	GQ127642	98	Soil
OS121	1 / 0 / 0 / 1	FJ584375	94	Acidic red soil
OS125	1 / 0 / 0 / 0	JN052746	99	Anaerobic digestion
OM66	0 / 0 / 0 / 1	AB600452	99	Iron-rich microbial mat

DISCUSSION

4.0 Environmental conditions

Surface seawater temperature observed in this study is typical of tropical coastal waters (Lee and Bong, 2008). Salinity was low at both KS and SM, but higher at PK and PD ($q > 5.64$, $p < 0.01$) (Table 3.0). As the samplings were all carried out during high tide levels, the lower salinity probably indicated a larger volume of freshwater that flowed into the KS and SM estuaries. The larger volume of river water would also explain the higher nitrate, ammonium and silicate concentrations observed at both KS and SM. The source of these nutrients is probably from anthropogenic activities and surface run-off around these sites (Bong and Lee, 2005).

Average TSS in all four sites ranged from 50 to 110 mg L⁻¹ in which POM constituted <20% (8.8 – 16.5 mg L⁻¹). This suggested that most of the TSS was inorganic in nature which reflected the prevalent conditions in these waters (Bong and Lee, 2005). Also, TSS increased the water turbidity, and caused poor irradiance, as reflected by the shallower Secchi disc depth. The average dissolved oxygen (DO) concentration at all four sites ranged between 140 – 210 µM. Both pH and DO concentration were lowest in SM compared to KS, PK and PD.

The waters at PD was characteristic of the most pristine condition with low inorganic nutrient concentrations and relatively high DO levels whereas PK waters had high nitrite concentrations and highest bacterial abundance. The conditions in PK might reflect some anthropogenic effects as the catchment area for Klang river includes the Klang valley which has the highest population density in Malaysia (Lee

and Bong, 2006). Although both bacterial abundance and Chl *a* concentration were within the range previously reported for various tropical ecosystems in this region (Lee et al., 2005, Lee et al., 2006, Lee and Bong, 2008) average Chl *a* concentration (Table 3.0) at KS was evidently higher than PK, PD and SM. The higher inorganic nutrient concentrations at KS supported a higher Chl *a* concentration but this was not observed at SM. Episodic nutrient inputs at estuaries may trigger phytoplankton production and subsequently an increase in Chl *a* concentration (Kennish, 1990). In contrast, the bacterial abundance was not significantly different among the sites.

4.1 Uncultured Archaea in different environments

Although studies have shown seasonal (Herfort et al., 2007) and temperature (Winter et al., 2008) influence on archaeal community, these were not considered in this study. The tropics experience a relatively stable temperature regime (Lee et al., 2009), with no seasonal change. Moreover, microbial community do not exhibit temporal variation in these waters (Lee et al., in preparation).

The observed numbers of OTUs in our clone libraries provided a minimum estimate of the total number of OTUs present. The number of uncollected OTUs present in the source environment is shown by the difference between the S_{Chao1} and S_{ACE} estimates with the observed OTUs (Aller and Kemp, 2008). Both estimators predicted that the source community in this study contains an estimated $48 \pm 22\%$ (from S_{Chao1}) and $139 \pm 105\%$ (from S_{ACE}) additional unique OTUs than were actually observed in the studied clone libraries. Since SM recorded highest diversity indices (Table 3.1), reflecting a phylogenetically diverse community, SM was probably more functionally diverse than KS, PK, and PD.

In this study, the phylogenetic analysis of 16S rDNA sequences showed high archaea diversity. Shannon index ranged from 2.16 – 3.24 (average = 2.60 ± 0.48), and was higher compared to Qinghai Lake (1.46 – 2.49) (Jiang et al., 2008) and East China Sea estuary (2.26 – 2.43) (Zeng et al., 2007). Our study also showed that archaea diversity was higher in KS and SM where the influence of river water was substantial, and this concurred with Galand et al. (2008). The number of OTUs obtained at $\geq 97\%$ similarity in this study was compared with the results from 28 other studies reviewed by Auguet et al. (2010) (Figure 4.0).

This cross-ecosystem approach seemed to suggest that the number of Archaea OTUs obtained decreased with increasing trophic states. Although more studies on Archaea population are required, studies of archaeal communities in sediments (Durbin and Teske, 2012) and lakes (Villaescusa et al., 2010) have also shown how both archaeal and bacterial diversity decreased with increasing trophic states.

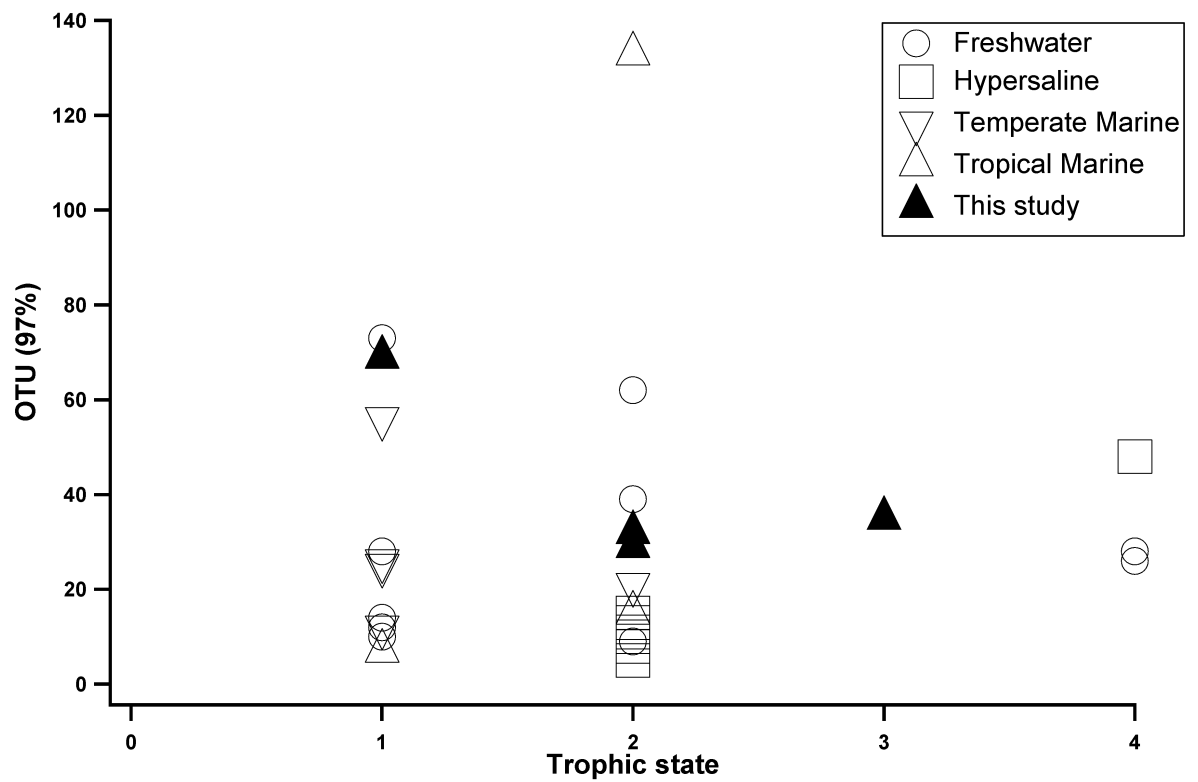


Figure 4.0 Diversity of uncultured archaea in studies from environments of different trophic state. Trophic state: 1 = oligotrophic, 2 = mesotrophic, 3 = eutrophic, 4 = hypertrophic. Trophic state is determined using Carlson's Trophic State Index (Carlson, 1977).

4.2 **Phylogenetic relatedness**

Most of the clones (> 50%) from SM and PD were affiliated to Euryarchaeota (Figure 3.6). Crenarchaeota clones represented a minor fraction (< 21%) in all four sites. The similarity of our sequences to sequences from various environments (Table 3.2) may indicate that these sequences are not limited by geographical barriers and are also present in tropical waters. The frequency of MCG clones was highest in SM but was represented by only a single clone in PD. Although inconclusive, the higher frequency of MCG clones in SM could be related to the comparatively higher ammonium, nitrate concentration, and lower salinity (Preston et al., 2011) in SM when compared to KS, PK and PD. We also found a high number of OTUs affiliated with Thaumarchaeota in KS. The ability of Thaumarchaeal populations to thrive in surface waters during eukaryotic phytoplankton blooms (Robidart et al. 2012) suggested that the higher Chl *a* concentration in KS could have played a role in Thaumarchaeota distribution.

However none of our sequences was identical to the reference sequences from cultured taxa, which is common in the 16S rRNA based surveys of natural samples (Fuhrman et al., 1992, McInerney et al., 1995, Kim et al., 2005). Some are found to be closely related to sequences from cultured methanogens previously isolated from diverse environments. Methanogenic clones observed in the SM clone library (6 clones), and KS clone library (2 clones) were related to the order *Methanosarcinales*. Only one clone was related (98%) to a methanogen in PK clone library whereas BLAST analysis did not show related methanogenic clones in the PD clone library.

Clone OS113 from KS was closely related (98%) to *Methanococcoides methylutens* which was isolated from marine sediments and particles of the Pacific Ocean (Cynar and Yayanos, 1991). Clone OM57 was similar (97%) to *Methanoplanus petrolearius* (Ollivier et al., 1997), isolated from petroleum wells, which was also observed in moderately polluted inner bay water in a tropical estuary (Vieira et al., 2007). A 99% similarity to *Methanocorpusculum labreanum*, isolated from lake surface sediment (Zhao et al., 1989) was exhibited by clone OM36. Sequences similar to this anaerobic methanogen were also retrieved from a eutrophic freshwater marsh (Chauhan et al., 2006). Although represented only by two clones (Clone OM58 and OM61), MBG-D group, commonly associated with methane-rich environments were present only in SM.

The physiological characteristics of an organism can be inferred from its phylogenetic position and the properties of their relatives. Thus, sequences that fall into a clade with known methanogens are likely to represent another methanogen in the environment (Robertson et al., 2005). Overall, there was a more diversified methanogenic archaea community at SM. The frequent occurrence of singletons in our clone libraries stress the necessity of more exhaustive sampling to obtain a fuller perspective on archaeal diversity (Robertson et al., 2005) but in contrast to bacterial libraries, a larger sampling size is not necessarily better when sampling Archaea (Aller and Kemp, 2008).

4.3 Techniques in evaluating archaeal diversity

A recent study which employed next-generation sequencing (NGS) techniques revealed a more detail archaeal communities, up to 700 OTUs from a polar mixed layer of the Arctic Ocean (Galand et al., 2009). Although the cloning method focuses on the dominant taxa (Zinger et al., 2012), Good's coverage showed that more than 85% of archaeal sequence types were obtained in this study. It is undeniable that we have uncovered via molecular ecological methods, an enormous richness in diversity. We were also able to observe several clades (Figure 3.9) that were not previously detected in tropical waters. The occurrence of unique sequences which does not cluster with sequences from the database is not an unusual phenomenon when studying environmental sequences (Bano et al., 2004, Huang et al., 2011, Bhattarai et al., 2012).

4.4 The expanding archaeal taxonomy

Although the deposition of archaeal sequences in the public databases is expanding dramatically, not all were affiliated with established archaeal groups, and this contributed to the complication of taxonomic assignments (Robertson et al., 2005). Previous introduction of synonyms and equivalent designations for the same archaea lineages has increased the confusion when assigning archaea sequences (Teske and Sorensen, 2008). The phylogenetic features of a given archaeal community are relatively conserved in its particular environment (Takai et al., 2001), and the presence of novel archaeal communities was evident after culture-independent analysis of rDNAs recovered from our tropical waters. However, more sequences from tropical waters are probably needed to support the phylogenetic placement of our sequences in the phylogenetic tree (Figure 3.9).

We did not assign specific names for these uncultured clusters to refrain from increasing the number of existing conflicting nomenclatures (DeSantis et al., 2006, Auguet et al., 2010). For example, the names of specific clusters such as SAGMCG-1 and MCG contradicted its phylogenetic placement since SAGMCG-1 were more likely to belong to the phylum Thaumarchaeota and MCG has yet to show clear affiliations to any established archaeal phyla (Pester et al., 2011). An Archaea phylogenetic review is therefore crucial (Robertson et al., 2005, Vieira et al., 2007). Although our study contributed sequences from tropical waters in our region to the developing database of environmental 16S rDNA clone libraries, more sequences are needed from various environments to support the placement of archaeal sequences to its respective clusters, and more importantly, to the right phylum.

4.5 Ordination analysis of OTUs to environmental variables

PCoA analyses showed that the archaeal communities of the four sites were clearly separated from each other where the first two principal coordinates (P1 and P2) explained 98.0 % of the total community variability. From the CCA analysis of archaeal OTUs in response to environmental variables (Figure 4.1), we also showed that the distribution of OTUs could be related to environmental variables.

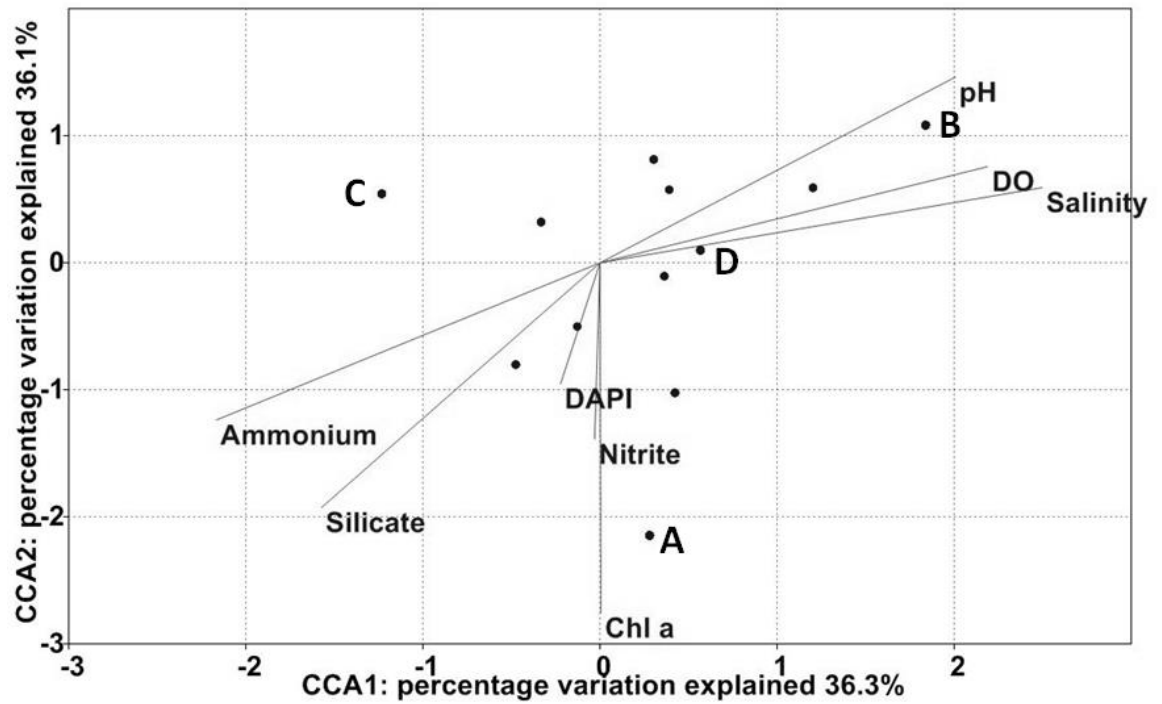


Figure 4.1 Correspondence canonical analysis (CCA) ordination plots for the first two dimensions of CCA of the relationship between the archaeal OTUs with environmental factors. OTUs present only in KS, PD and SM are indicated by points A, B and C, respectively. OTUs present only in PK was indicated by point D.

This plot shows the apparent collinearity of some of the environmental variables such as pH, dissolved oxygen concentration and salinity, that reflected freshwater influence. However ammonium showed an inverse relationship to variables related to freshwater inflow. Freshwater inflow provided nutrients and as freshwater inflow increased, salinity decreased. The first two axes of the CCA (CCA1 and CCA2) of the archaeal OTUs versus environmental variables explained 72.4% of the cumulative variance of the OTU-environment relationship (Figure 4.1). Our CCA model illustrated that Chl *a* concentration contributed significantly ($P < 0.001$, 1,000 Monte Carlo permutations) to the OTU-environment relationship, explaining 36.1% of the total variance. OTUs present only in KS, PD and SM were indicated by point A, B and C, respectively (Figure 4.1). OTUs present only in PK was indicated by point D.

From our CCA model, we observed two distinct groups that were placed away from the center. Point A was comprised of 21 OTUs whereby 10 OTUs belonged to the MCG and 7 OTUs were affiliated with Thaumarchaeota. Two other OTUs similar ($> 97\%$) to the methanogens; *Methanococcoides methylutens* and *Methanosaeta concilii* were also found to be present only at KS. Although affiliated with Euryarchaea, the two remaining unique OTUs at KS showed a rather low similarity with sequences from the GenBank database ($< 95\%$). Concurrent with the domination of Euryarchaea at PD, 16 out of 20 unique OTUs at PD (point B) belonged to the phylum Euryarchaeota. Although Point A consisted of mainly OTUs belonging to MCG (10 OTUs), only 1 OTU at point B were affiliated with MCG sequence previously retrieved from mangrove soil and 3 OTUs were affiliated with Thaumarchaeota (Table 3.2). The archaeal OTUs present only in KS reflected the effects of higher Chl *a* concentration whereas archaeal OTUs present only at PD

mirrored the seawater influence. This shows that although archaeal marine groups seemed to be widely distributed (Vieira et al. 2007), archaeal OTUs are still restricted to a particular geographical or environmental conditions.

CONCLUSION

This study investigated the archaeal communities via 16S rRNA clonal library in the tropical waters of Malaysia, and found that the archaeal community differed significantly among the different sites. The differences among the archaeal community were related to environmental variables. There were also a large number of sequences that did not affiliate with known groups, and showed a need for more extensive sampling.

REFERENCES

- Ahmad N, Johri S, Sultan P, Abdin MZ, Qazi GN. 2011. Phylogenetic characterization of archaea in saltpan sediments. *Indian J. Microbiol.* 51: 132-137.
- Aller JY, Kemp PF. 2008. Are Archaea inherently less diverse than Bacteria in the same environments? *FEMS Microbiol. Ecol.* 65: 74-87.
- Alquéres SMC, Almeida RV, Clementino MM, Vieira RP, Almeida WI, Cardoso AM, Martins OB. 2007. Exploring the biotechnological applications in the Archaeal domain. *Braz. J. Microbiol.* 38: 398-405.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Auchtung TA, Christina D, Vesbach T, Cavanaugh CM. 2006. 16S rRNA Phylogenetic investigation of the Candidate division “Korarchaeota”. *Appl. Environ. Microbiol.* 72: 5066-5082.
- Auchtung TA, Shyndriayeva G, Cavanaugh CM. 2011. 16S rRNA phylogenetic analysis and quantification of Korarchaeota indigenous to the hot springs of Kamchatka, Russia. *Extremophiles* 15: 105–116.
- Auguet JC, Barberan A, Casamayor EO. 2010. Global ecological patterns in uncultured Archaea. *ISME J.* 4: 182 – 190.
- Bano N, Ruffin S, Ransom B, Hollibaugh JT. 2004. Phylogenetic composition of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages. *Appl. Environ. Microbiol.* 70: 781-789.
- Barns SM, Delwiche CF, Palmer JD, Pace NR. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 93: 9188-9193.

- Bergmann I, Nettmann E, Mundt K, Klocke M. 2010. Determination of methanogenic Archaea abundance in a mesophilic biogas plant based on 16S rRNA gene sequence analysis. *Can. J. Microbiol.* 56: 440-444.
- Bhattarai S, Ross KA, Schmid M, Anselmetti FS, Bürgmann H. 2012. Local conditions structure unique archaeal communities in the anoxic sediments of meromictic Lake Kivu. *Microb. Ecol.* 64: 291–310.
- Bolhuis H, Stahl LJ. 2011. Analysis of bacterial and archaeal diversity in coastal microbial mats using massive parallel 16S rRNA gene tag sequencing. *ISME J.* 11: 1701–1712.
- Bomberg M, Jurgens G, Saano A, Sen R, Timonen S. 2003. Nested PCR detection of Archaea in defined compartments of pine mycorrhizospheres developed in boreal forest humus microcosms. *FEMS Microbiol. Ecol.* 43: 163-171.
- Bong CW, Lee CW. 2005. Microbial abundance and nutrient concentration in riverine and coastal waters of North-East Langkawi. *Malays. J. Sci.* 24: 29–35.
- Boone DR, Castenholz RW. 2001. *Bergey's Manual of Systematic Bacteriology*. Springer; New York.
- Bostrom KH, Simu K, Hagstrom A, Riemann L. 2004. Optimization of DNA extraction for quantitative marine bacterioplankton community analysis. *Limnol. Oceanogr.* 2: 365–373.
- Braker G, Rio H, Devol A, Fesefeldt A, Tiedje J. 2001. Community structure of denitrifiers, bacteria, and archaea along redox gradients in pacific northwest marine sediments by terminal restriction fragment length polymorphism analysis of amplified nitrite reductase (nirS) and 16S rRNA genes. *Appl. Environ. Microbiol.* 67: 1893-1901.
- Brochier-Armanet C, Boussau B, Gribalso S, Forterre P. 2008. Mesophilic

- Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6: 245-252.
- Buckley DH, Graber JR, Schmidt TM. 1998. Phylogenetic analysis of nonthermophilic members of the kingdom crenarchaeota and their diversity and abundance in soils. *Appl. Environ. Microbiol.* 64: 4333-4339.
- Burggraf S, Mayer T, Amann R, Schadhauer S, Woese CR, Stetter KO. 1994. Identifying members of the domain Archaea with rRNA-targeted oligonucleotide probes. *Appl. Environ. Microbiol.* 60: 3112-3119.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7: 335-336.
- Carlson RE. 1977. A trophic state index for lakes. *Limnology and Oceanography.* 22: 361-369.
- Casanueva A, Galada N, Baker GC, Grant WD, Heaphy S, Jones B, Yanhe M, Ventosa A, Blamey J, Cowan DA. 2008. Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. *Extremophiles* 12: 651–656.
- Chauhan A, Reddy KR, Ogram AV. 2006. Syntrophic-archaeal associations in a nutrient-impacted freshwater marsh. *J. Appl. Microbiol.* 100: 73-84.
- Chao A, Ma M-C, Yang MCK. 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. *Biometrika* 80: 193–201.
- Church MJ, DeLong EF, Ducklow HW, Karner MB, Preston CM, Karl DM. 2003.

- Abundance and distribution of planktonic Archaea and Bacteria in the waters west of the Antarctic Peninsula. *Limnol. Oceanogr.* 48: 1893-1902.
- Climate Avenue, 2011. Mangrove forests, ecotourism and their services to mankind. Retrieved September 1, 2011 from the World Wide Web: <http://www.climateavenue.com/forest.mangr.ecotourism.htm>
- Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM. 2007. The ribosomal database project (RDP-II): Introducing myRDP space and quality controlled public data. *Nucleic Acids Res.* 35: 169–172.
- Collins RE, Rocap G, Deming JW. 2010. Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Env. Microbiol.* 12: 1828-1841.
- Cynar FJ, Yayanos AA. 1991. Enrichment and characterization of a methanogenic bacterium from the upper layer of the ocean. *Curr. Microbiol.* 23: 89-96.
- Cytryn E, Minz D, Oremland RS, Cohen Y. 2000. Distribution and diversity of Archaea corresponding to the limnological cycle of a hypersaline stratified lake (Solar Lake, Sinai, Egypt). *Appl. Environ. Microbiol.* 66: 3269–3276.
- Dalgaard JZ, Garrett RA. 1993. Archaeal hyperthermophile genes. In M. Kates, D.J. Kushner, and A.T. Matheson (ed.), *The Biochemistry of Archaea (Archaeobacteria)*. Elsevier, Amsterdam, The Netherlands.
- DeLong EF. 1992. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. U.S.A.* 89: 5685-5689.
- DeLong EF. 1998. Everything in moderation: archaea as 'non-extremophiles'. *Curr. Opin. Genet. Devel.* 8: 649–654.
- DeLong EF. 2003. Oceans of Archaea. *ASM News* 69: 503-511.
- DeLong EF, Taylor LT, Marsh TL, Preston CM. 1999. Visualization and enumeration

- of marine planktonic Archaea and Bacteria by using polyribonucleotide probes and fluorescent in situ hybridization. *Appl. Environ. Microbiol.* 12: 5554-5563.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72: 5069–5072.
- Diez B, Pedros-alio C, Marsh TL, Massana R. 2001. Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and comparison of DGGE with other molecular techniques. *Appl. Environ. Microbiol.* 67: 2942-2951.
- Dunbar J, Takala S, Barns SM, Davis JA, Kuske CR. 1999. Levels of bacterial community diversity in four arid soils compared by cultivation and 16S rRNA gene cloning. *Appl. Environ. Microbiol.* 65: 1662-1669.
- Durbin AM, Teske A. 2012. Archaea in organic-lean and organic-rich marine subsurface sediments: an environmental gradient reflected in distinct phylogenetic lineages. *Front. Microbiol.* 3: 1-26.
- Edgcomb VP, Molyneux SJ, Böer S, Wirsén CO, Saito M, Atkins MS, Lloyd K, Teske A. 2007. Survival and growth of two heterotrophic hydrothermal vent archaea, *Pyrococcus* strain GB-D and *Thermococcus fumicolans*, under low pH and high sulfide concentrations in combination with high temperature and pressure regimes. *Extremophiles* 11: 329–342.
- Einen J, Thorseth IH. 2008. Enumeration of Archaea and Bacteria in seafloor basalt using real-time quantitative PCR and fluorescence microscopy. *FEMS Microbiol. Lett.* 282: 182-187.
- Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, Hedlund BP,

- Brochier-Armanet C, Kunin V, Anderson I, Lapidus A, Goltsman E, Barry K, Koonin EV, Hugenholtz P, Kyrpides N, Wanner G, Richardson P, Keller M, Stetter KO. 2008. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 105: 8102–8107.
- Elshahed MS, Najar FZ, Roe BA, Oren A, Dewers TA, Krumholz LR. 2004. Survey of Archaeal diversity reveals an abundance of halophilic Archaea in a low-salt, sulfide- and sulfur-rich spring. *Appl. Environ. Microbiol.* 70: 2230-2239.
- Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. 2005. Ubiquity and diversity of ammonia oxidizing archaea in water columns and sediments of the ocean. *Proc. Natl. Acad. Sci. U.S.A.* 102: 14683–14688.
- Fuhrman JA, McCallum K, Davis AA. 1992. Novel major archaebacterial group from marine plankton. *Nature* 356: 148-149.
- Fuhrman JA, Ouverney CC. 1998. Marine microbial diversity studied via 16S rRNA sequences: cloning results from coastal waters and counting of native archaea with fluorescent single cell probes. *Aquat. Ecol.* 32: 3-15.
- Galand PE, Lovejoy C, Vincent WF. 2006. Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquat. Microb. Ecol.* 44: 115-126.
- Galand PE, Lovejoy C, Pouliot J, Vincent WF. 2008. Heterogeneous archaeal communities in the particle-rich environment of an arctic shelf ecosystem. *J. Mar. Syst.* 74: 774-782.
- Galand PE, Casamayor EO, Kirchman DL, Potvin M, Lovejoy C. 2009. Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J.* 3: 860-869.
- Goeyens L, Kindermans N, Abu Yusuf M, Elskens M. 1998. A room temperature

- procedure for the manual determination of urea in seawater. *Estuar. Coast. Shelf Sci.* 47: 415-418.
- Goh F, Leuko S, Allen M, Bowman J, Kamekura M, Neilan B, Burns B. 2006. *Halococcus hamelinensis* sp. nov., a novel halophilic archaeon isolated from stromatolites in Shark Bay, Australia. *Int. J. Syst. Evol. Microbiol.* 56: 1323-1329.
- Gómez-Silván C, Molina-Muñoz M, Poyatos JM, Ramos A, Hontoria E, Rodelas B, González L. 2010. Structure of archaeal communities in membrane-bioreactor and submerged-biofilter wastewater treatment plants. *Bioresour. Technol.* 101: 2096-2105.
- Good IJ. 1953 The population frequencies of species and the estimation of population parameters. *Biometrika* 40: 237–264.
- Grasshoff K, Kremling K, Ehrhardt, M. 1999. *Methods of seawater analysis*, 3rd edition. Wiley-VCH, Weinheim, 600 pp.
- Gribaldo S, Brochier-Armanet C. 2006. The origin and evolution of Archaea: a state of the art. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361: 1007-1022.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis. <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>.
- Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J, Preston C, de la Torre J, Richardson PM, DeLong EF. 2006. Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc. Natl. Acad. Sci. U.S.A.* 103: 18296–18301.
- Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*. 4: 9pp.http://palaeo-electronica.org/2001_1/past/issue1_01.htm.
- Hatzenpichler R, Lebedeva EV, Spieck E, Stoecker K, Richter A, Daims H,

- Wagner M. 2008. A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc. Natl. Acad. Sci. U.S.A.* 105: 2134–2139.
- Herfort L, Schouten S, Abbas B, Veldhuis MJW, Coolen MJL, Wuchter C, Boon JP, GJ, Damste JSS. 2007. Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiol. Ecol.* 62: 242-257.
- Herfort L, Kim JH, Coolen MJL, Abbas B, Schouten S, Herndl GJ, Damste JS. 2009. Diversity of Archaea and detection of crenarchaeotal *amoA* genes in the rivers Rhine and Têt. *Aquat. Microb. Ecol.* 55: 189–20.
- Hohn MJ, Hedlund BP, Huber H. 2002. Detection of 16S rDNA sequences representing the novel phylum “Nanoarchaeota”: indication for a broad distribution in high temperature. *Syst. Appl. Microbiol.* 25: 551–554.
- Huang Q, Dong CZ, Dong RM, Jiang H, Wang S, Wang G, Fang B, Ding X, Niu L, Li X, Zhang C, Dong H. 2011. Archaeal and bacterial diversity in hot springs on the Tibetan Plateau, China. *Extremophiles* 15: 549-63.
- Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417: 63-67.
- Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K., Nealson KH, Horikoshi K. 2003. Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Appl. Environ. Microbiol.* 69: 7224–7235.
- Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever MA, Lauer A, Suzuki M, Takai K, Delwiche M, Colwell FS, Nealson KH, Horikoshi K, D’Hondt S, Jorgensen BB. 2006. Biogeographical distribution and diversity of microbes

- in methane hydrate-bearing deep marine sediments on the Pacific ocean margin. *Proc. Natl. Acad. Sci. U.S.A.* 103: 2815–2820.
- Hashim Z. 2010. Development of atmospheric based flood forecasting and warning system selected for river basins in Malaysia. In: National Seminar on Meterology: “Enhancing Weather and Climate Services”, 9 Nov 2010, Kuala Lumpur.
- Jiang H, Dong H, Yu B, Ye Q, Shen J, Rowe H, Zhang C. Dominance of putative marine benthic Archaea in Qinghai Lake, north-western China. 2008. *Env. Microbiol.* 10: 2355-2367.
- Juottonen H, Tuittila E, Juutinen S, Fritze H, Yrjala K. 2008. Seasonality of rDNA- and rRNA-derived archaeal communities and methanogenic potential in a boreal mire. *ISME J.* 2: 1157-1168.
- Karner MB, DeLong EF, Karl DM. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409: 507-510.
- Kellogg CA. 2004. Tropical Archaea: diversity associated with the surface microlayer of corals. *Mar. Ecol. Prog. Ser.* 272: 81-88.
- Kennish MJ. 1990. ‘Ecology of Estuaries, vol II: Biological Aspects.’ CRC Press: Boca Raton.
- Kim BS, Oh HM, Kang H, Chun J. 2005. Archaeal diversity in tidal flat sediment as revealed by 16S rDNA analysis. *J. Microbiol.* 43: 144-151.
- Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437: 543–546.
- Kotchoni SO, Gachomo EW, Betiku E, Shonukan OO. 2003. A home-made kit for plasmid DNA mini-preparation. *Afr. J. Biotechnol.* 2: 88-90.
- Lam P, Jensen MM, Lavik G, McGinnis DF, Muller B, Schubert CJ, Amann R,

- Thamdrup B, Kuypers MMM. 2007. Linking crenarchaeal and bacterial nitrification to anammox in the Black Sea. *Proc. Natl. Acad. Sci. U.S.A.* 104: 7104-7109.
- Law AT, Kuan CY, Jan JK, Hong S. 2002. Hydrocarbon and sewage pollution in the coastal waters of Port Dickson, Straits of Malacca. In: *Tropical Marine Environment: Charting Strategies for the Millenium*, Yosuff FM, Ibrahim KM, Tan SG, Tai SY (Eds.). M. Shariff Serdang Publisher, Malaysia.
- Lee CW, Bong CW, Muhammad Yusof MA, Alias SA. 2005. Bacteria-mediated carbon flux in mangrove waters: A Malaysian perspective. *Int. J. Ecol. Environ. Sci.* 31: 203-211.
- Lee CW, Bong CW. 2006. Carbon flux through bacteria in a eutrophic tropical environment: Port Klang waters. In: Wolanski E (ed.) *The Environment in Asia Pacific Harbours*. Springer, The Netherlands: 329–345.
- Lee CW, Bong CW, Ng CC, Alias SA. 2006. Factors affecting variability of heterotrophic and phototrophic microorganisms at high water in a mangrove forest at Cape Rachado, Malaysia. *Malays. J. Sci.* 25: 55–66.
- Lee CW, Bong CW, Yii SH. 2009. Temporal variation of bacterial respiration and growth efficiency in tropical coastal waters. *Appl. Environ. Microbiol.* 75: 7594-7601.
- Lee CW, Bong CW. 2008. Bacterial abundance and production, and their relation to primary production in tropical coastal waters of Peninsular Malaysia. *Mar. Freshwater Res.* 59: 10-21.
- Leininger S, Urich T, Schlöter M, Schwark L, Qi J, Nicol GW, Porsser JI, Schuster SC, Schleper C. 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442: 806-809.

- Lliros M, Gich F, Plascencia A, Auguet J, Darchambeau F, Casamayor EO, Descy J, Borrego C. 2010. Vertical distribution of ammonia-oxidizing crenarchaeota and methanogens in the epipelagic waters of Lake Kivu. *Appl. Environ. Microbiol.* 76: 6853-6863.
- Longhurst AR., Pauly D. 1987. 'Ecology of Tropical Oceans.' Academic Press: San Diego, CA.
- Lopez-Garcia P, Moreira D, Lopez-lopez A, Rodriguez-Valera F. 2001. A novel haloarchaeal-related lineage is widely distributed in deep oceanic regions. *Env. Microbiol.* 3: 72-78.
- Luna GM, Stumm K, Pusceddu A, Danovaro R. 2009. Archaeal diversity in deep-sea sediments estimated by means of different terminal-restriction length polymorphisms (T-RFLP) protocols. *Curr. Microbiol.* 59: 356-361.
- Margalef R. 1958. Information theory in ecology. *Gen. Syst.* 3:36–71.
- Margesin R, Miteva V. 2011. Diversity and ecology of psychrophilic microorganisms. *Res. Microbiol.* 3: 346-361.
- Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA. 2009. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* 461: 976-979.
- Massana R, Murray AE, Preston CM, DeLong EF. 1997. Vertical distribution and phylogenetic characterization of marine planktonic archaea in the Santa Barbara Channel. *Appl. Environ. Microbiol.* 63: 50–56.
- McInerney JO, Wilkinson M, Patching JW, Embley TM, Powell R. 1995. Recovery and phylogenetic analysis of novel archaeal rRNA sequences from a deep-sea deposit feeder. *Appl. Environ. Microbiol.* 61: 1646-1648.
- Miller-Coleman RL, Dodsworth JA, Ross CA, Shock EL, Williams AJ, Hartnee HE, McDonald AI, Javig JR, Hedlund BP. 2012. Korarchaeota diversity,

- biogeography, and abundance in Yellowstone and Great Basin Hot Springs and ecological niche modeling Based on machine learning. PLoS ONE 7(5): e35964. doi:10.1371/journal.pone.0035964.
- Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF. 2007. Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. Environ. Microbiol. 9: 1162–1175.
- Moyer CL, Tiedje JM, Dobbs FC, Karl DM. 1998. Diversity of deep-sea hydrothermal vent Archaea from Loihi Seamount, Hawaii. Deep-Sea Res. Part II 45: 303-317.
- Murray AE, Preston CM, Massana R, Taylor LT, Blakis A, Wu K, DeLong EF. 1998. Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. Appl. Environ. Microbiol. 64: 2585–2595.
- Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee GJ, Hattori M, Kanai A, Atomi H, Takai K, Takami H. 2011. Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res. 39: 3204–3223.
- Ollivier B, Cayol JL, Patel BKC, Magot M, Fardeau ML, Garcia JL. 1997. *Methanoplanus petrolearius* sp. nov., a novel methanogenic bacterium from an oil-producing well. FEMS Microbiol. Lett. 147: 51-56.
- Parsons TR, Maita Y, Lalli CM. 1984. A manual of chemical and biological methods for seawater analysis. Pergamon Press, Oxford, 173 pp.
- Perevalova AA, Kolganova TV, Birkeland N, Schleper C, Osmolovskaya EA,

- Lebedinsky A. 2008. Distribution of Crenarchaeota representatives in terrestrial hot springs of Russia and Iceland. *Appl. Environ. Microbiol.* 74: 7620-7628.
- Pester M, Schleper C, Wagner M. 2011. The Thaumarchaeota: an emerging view of their phylogeny and Ecophysiology. *Curr. Opin. Microbiol.* 14: 300–306.
- Preston CM, Harris A, Ryan JP, Roman B, Marin III R, Jensen S, Everlove C, Birch J, Dzenitis JM, Pargett D, Adachi M, Turk K, Zehr JP, Scholin CA. 2011. Underwater application of quantitative PCR on an ocean mooring. *PLoS ONE* 6(8): e22522. doi:10.1371/journal.pone.0022522.
- Poplawski AB, Mårtensson L, Wartainen I, Rasmussen U. 2007. Archaeal diversity and community structure in a Swedish barley field: Specificity of the EK510R/(EURY498) 16S rDNA primer. *J. Microbiol. Methods.* 69: 161-173.
- Porter KG, Feig YS. 1980. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* 25: 943–948.
- Port Klang Malaysia, 2011. Port Klang Authority's Official Portal. Retrieved October 5, 2011 from the World Wide Web: <http://www.pka.gov.my/phocadownload/gateway/gateway%202011.pdf>
- Ramírez-Flandes S, Ulloa O. 2008. Bosque: Integrated phylogenetic analysis software. *Bioinformatics* 24: 2539-2541. doi: 10.1093/bioinformatics/btn466.
- Reigstad LJ, Jorgensen SL, Schleper C. 2010. Diversity and abundance of Korarchaeota in terrestrial hot springs of Iceland and Kamchatka. *ISME J.* 4: 346-56.
- Robertson CE, Harris JK, Spear JR, Pace NR. 2005. Phylogenetic diversity and ecology of environmental Archaea. *Curr. Opin. Microbiol.* 8: 638–642.
- Robidart JC, Preston CM, Paerl RW, Turk KA, Mosier AC, Francis CA, Scholin CA,

- Zehr JP. 2012. Seasonal *Synechococcus* and *Thaumarchaeal* population dynamics examined with high resolution with remote in situ instrumentation. *ISME J.* 6: 513-523.
- Roh SW, Kim K, Nam Y, Chang H, Park E, Bae J. 2010. Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *ISME J.* 4: 1–16.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75: 7537- 7541.
- Siboni N, Ben-Dov E, Sivan A, Kushmaro A. 2008. Global distribution and diversity of coral-associated Archaea and their possible role in the coral holobiont nitrogen cycle. *Environ. Microbiol.* 10: 2979–2990.
- Spang A, Hatzenpichler R, Brochier-Armanet C, Rattei T, Tischler P, Spieck E, Streit W, Stahl DA, Wagner M, Schleper C. 2010. Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum *Thaumarchaeota*. *Trends Microbiol.* 18: 331-340.
- Stoica E. 2009. Application of T-RFLP analysis to the study of the coastal Black Sea. *Rom. Biotechnol. Lett.* 14: 4710-4719.
- Stoica E, Herndl GJ. 2007. Contribution of Crenarchaeota and Euryarchaeota to the prokaryotic plankton in the coastal northwestern Black Sea. *J. Plankton Res.* 29: 699-706.
- Takai K, Horikoshi K. 1999. Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* 152: 1285–1297.
- Takai, K, Moser DP, DeFlaun M, Onstott TC, Fredrickson JK. 2001. Archaeal

- diversity in waters from deep south African Gold Mines. *Appl. Environ. Microbiol.* 67: 5750–5760.
- Teske A, Hinrichs KU, Edgcomb V, Gomez AdV, Kysela D, Sylva SP, Sogin ML, Jannasch HW. 2002. Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Appl. Environ. Microbiol.* 68: 1994–2007.
- Teske A, Sørensen KB. 2008. Uncultured archaea in deep marine subsurface sediments: have we caught them all? *ISME J.* 2: 3-18.
- Vetriani C, Jannasch HW, MacGregor BJ, Stahl DA, Reysenbach AL. 1999. Population structure and phylogenetic characterization of marine benthic archaea in deep-sea sediments. *Appl. Environ. Microbiol.* 65: 4375–4384.
- Vieira RP, Clementino MM, Cardoso AM, Oliveira DN, Albano RM, Gonzalez AM, Paranhos R, Martins OB. 2007. Archaeal communities in a tropical estuarine ecosystem: Guanabara Bay, Brazil. *Microb. Ecol.* 54: 460-468.
- Villaescusa JA, Casamayor EO, Rochera C, Velazquez D, Chicote A, Quesada A, Camacho A. 2010. A close link between bacterial community composition and environmental heterogeneity in maritime Antarctic lakes. *Int. Microbiol.* 13: 67-77.
- Weidler GW, Gerbl FW, Stan-Lotter H. 2008. Crenarchaeota and their role in the nitrogen cycle in a subsurface radioactive thermal spring in the Austrian Central Alps. *Appl. Environ. Microbiol.* 74:5934-5942.
- Winter C, Moeseneder MM, Herndl GJ, Weinbauer MG. 2008. Relationship of geographic distance, depth, temperature, and viruses with prokaryotic communities in the eastern tropical Atlantic Ocean. *Microb. Ecol.* 56: 383-389.

- Wright ES, Yilmaz LS, Noguera DR. 2011. DECIPHER: A search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* doi:10.1128/AEM.06516-11.
- Wright JJ, Konwar KM, Hallam SJ. 2012. Microbial ecology of expanding oxygen minimum zones. *Nat. Rev.* 10: 381-394.
- Wuchter C, Abbas B, Coolen MJL, Herfort L, Bleijswijk JV, Timmers P, Strous M, Teira E, Herndl GJ, Middelburg JJ, Schouten S, Damste JSS. 2006. Archaeal nitrification in the ocean. *Proc. Natl. Acad. Sci. U.S.A.* 33: 12317-12322.
- Yan B, Hong K, Yu ZN. 2006. Archaeal communities in mangrove soil characterized by 16S rRNA gene clones. *J. Microbiol.* 44: 566-571.
- Yap CK, Ismail A, Tan SG, Omar H. 2002. Concentrations of Cu and Pb in the offshore and intertidal sediments of the west coast of Peninsular Malaysia. *Environ. Int.* 28: 467-479.
- Zeng Y, Li H, Jiao N. 2007. Phylogenetic diversity of planktonic archaea in the estuarine region of East China Sea. *Microbiol. Res.* 162: 26 – 36.
- Zhang W, Sørensen G, Li T, Yu X, Zhang L. 2010. Diversity and community structure of Archaea in deep subsurface sediments from the tropical western Pacific. *Curr. Microbiol.* 60: 439-445.
- Zhao Y, Boone DR, Mah RA, Boone JE, Xun L. 1989. Isolation and characterization of *Methanocorpusculum labreanum* sp. nov. from the LaBrea Tar Pits. *Int. J. Syst. Bacteriol.* 39: 10-13.
- Zinger L, Gobet A, Pommier T. 2012. Two decades of describing the unseen majority of aquatic microbial diversity. *Mol. Ecol.* 21: 1878–1896.

Appendix A

Preparation of the reagents used in determination of urea concentration.

Diacetylmonoxime solution

Diacetylmonoxime	8.5 g
Distilled water	250 ml

Thiosemicarbazide solution

Thiosemicarbazide	0.95 g
Distilled water	100 ml

To prepare Reagent A:

1. Thiosemicarbazide solution is prepared.
2. Diacetylmonoxime solution is prepared.
3. 10 mL of thiosemicarbazide solution is added into 250 mL diacetylmonoxime solution to obtain reagent A.

Ferric chloride solution

Ferric chloride	0.15 g
dH ₂ O	10 ml

To prepare Reagent B:

1. Ferric chloride solution is prepared.

2. 300 ml concentrated H_2SO_4 is diluted to 535 ml in distilled water. Sulphuric acid solution is then allowed to cool.
3. 0.5 mL ferric chloride solution is then added to sulphuric acid solution.

Appendix B

Sequences of unique OTU were deposited in the GenBank database under the following accession numbers:

Affiliation	OTU	Accession	Description
Euryarchaeota	OS113	JQ 626967	Uncultured archaeon clone KS_114 16S ribosomal RNA gene, partial sequence
	OS128	JQ 626961	Uncultured archaeon clone KS_219 16S ribosomal RNA gene, partial sequence
	OM80	JQ 626958	Uncultured archaeon clone SM_26 16S ribosomal RNA gene, partial sequence
	OM57	JQ 626887	Uncultured archaeon clone SM_295 16S ribosomal RNA gene, partial sequence
	OM60	JQ 626892	Uncultured archaeon clone SM_27 16S ribosomal RNA gene, partial sequence
	OM49	JQ 626878	Uncultured archaeon clone SM_42 16S ribosomal RNA gene, partial sequence
	OM9	JQ 626928	Uncultured archaeon clone PK_147 16S ribosomal RNA gene, partial sequence
	OM36	JQ 626980	Uncultured archaeon clone SM_24 16S ribosomal RNA gene, partial sequence
	OK102	JQ 627016	Uncultured archaeon clone PK_291 16S ribosomal RNA gene, partial sequence
	OD91	JQ 626873	Uncultured archaeon clone PD_311 16S ribosomal RNA gene, partial sequence
	OM5	JQ 626889	Uncultured archaeon clone SM_98 16S ribosomal RNA gene, partial sequence
	OD94	JQ 626943	Uncultured archaeon clone PD_92 16S ribosomal RNA gene, partial sequence
	OD25	JQ 626940	Uncultured archaeon clone PD_125 16S ribosomal RNA gene, partial sequence
	OM2	JQ 626834	Uncultured archaeon clone PD_297 16S ribosomal RNA gene, partial sequence
	OM24	JX103592	Uncultured archaeon clone PD_349 16S ribosomal RNA gene, partial sequence
	OM3	JX 103572	Uncultured archaeon clone SM_109R 16S ribosomal RNA gene, partial sequence
	OD92	JQ 626875	Uncultured archaeon clone PD_163 16S ribosomal RNA gene, partial sequence
	OD21	JQ 626845	Uncultured archaeon clone PD_43 16S ribosomal RNA gene, partial sequence
	OK22	JQ 626844	Uncultured archaeon clone PK_140 16S ribosomal RNA gene, partial sequence
	OM71	JQ 626842	Uncultured archaeon clone SM_163 16S ribosomal RNA gene, partial sequence
	OD82	JQ 627014	Uncultured archaeon clone PD_174 16S ribosomal RNA gene, partial sequence
	OM27	JQ 626985	Uncultured archaeon clone SM_217 16S ribosomal RNA gene, partial sequence
	OD28	JQ 626949	Uncultured archaeon clone PD_187 16S ribosomal RNA gene, partial sequence
	OD88	JX 103583	Uncultured archaeon clone PD_357R 16S ribosomal RNA gene, partial sequence
	OD96	JQ 626947	Uncultured archaeon clone PD_71 16S ribosomal RNA gene, partial sequence

	OD85	JQ 627018	Uncultured archaeon clone PD_167 16S ribosomal RNA gene, partial sequence
	OM20	JQ 626862	Uncultured archaeon clone SM_148 16S ribosomal RNA gene, partial sequence
Euryarchaeota	OD89	JX 103589	Uncultured archaeon clone PD_18 16S ribosomal RNA gene, partial sequence
	OD95	JQ 626945	Uncultured archaeon clone PD_265 16S ribosomal RNA gene, partial sequence
	OD90	JX 103590	Uncultured archaeon clone PD_221 16S ribosomal RNA gene, partial sequence
	OD83	JX 103581	Uncultured archaeon clone PD_178R 16S ribosomal RNA gene, partial sequence
	OD23	JQ 626938	Uncultured archaeon clone PD_195 16S ribosomal RNA gene, partial sequence
	OM31	JQ 626867	Uncultured archaeon clone PK_301 16S ribosomal RNA gene, partial sequence
	OM32	JQ 626865	Uncultured archaeon clone SM_318 16S ribosomal RNA gene, partial sequence
	OD30	JQ 626870	Uncultured archaeon clone PK_173 16S ribosomal RNA gene, partial sequence
	OM29	JQ 626959	Uncultured archaeon clone SM_268 16S ribosomal RNA gene, partial sequence
	OM4	JQ 626849	Uncultured archaeon clone SM_319 16S ribosomal RNA gene, partial sequence
	OM47	JQ 626874	Uncultured archaeon clone SM_165 16S ribosomal RNA gene, partial sequence
	OK103	JX 103578	Uncultured archaeon clone PK_167R 16S ribosomal RNA gene, partial sequence
	OM1	JQ 626937	Uncultured archaeon clone PD_158 16S ribosomal RNA gene, partial sequence
	OD98	JQ 626956	Uncultured archaeon clone PD_76 16S ribosomal RNA gene, partial sequence
	OM48	JQ 626877	Uncultured archaeon clone SM_218 16S ribosomal RNA gene, partial sequence
	OM68	JQ 626900	Uncultured archaeon clone SM_74 16S ribosomal RNA gene, partial sequence
	OM43	JQ 627011	Uncultured archaeon clone SM_191R 16S ribosomal RNA gene, partial sequence
	OD87	JQ 626868	Uncultured archaeon clone PD_123 16S ribosomal RNA gene, partial sequence
	OK106	JQ 626926	Uncultured archaeon clone PK_289 16S ribosomal RNA gene, partial sequence
	OM61	JQ 626893	Uncultured archaeon clone SM_1 16S ribosomal RNA gene, partial sequence
	OM41	JQ 627004	Uncultured archaeon clone SM_34 16S ribosomal RNA gene, partial sequence
	OM75	JQ 626910	Uncultured archaeon clone SM_302 16S ribosomal RNA gene, partial sequence
	OM70	JQ 626906	Uncultured archaeon clone SM_121 16S ribosomal RNA gene, partial sequence
	OM55	JQ 626884	Uncultured archaeon clone SM_267 16S ribosomal RNA gene, partial sequence
	OM53	JQ 626882	Uncultured archaeon clone SM_237 16S ribosomal RNA gene, partial sequence
	OM46	JQ 626869	Uncultured archaeon clone SM_162 16S ribosomal RNA gene, partial sequence
	OM63	JQ 626895	Uncultured archaeon clone SM_199 16S ribosomal RNA gene, partial sequence

	OM52	JQ 626881	Uncultured archaeon clone SM_275 16S ribosomal RNA gene, partial sequence
	OM58	JQ 626890	Uncultured archaeon clone SM_69 16S ribosomal RNA gene, partial sequence
	OM79	JQ 626953	Uncultured archaeon clone SM_151R 16S ribosomal RNA gene, partial sequence
	OM59	JQ 626891	Uncultured archaeon clone SM_46 16S ribosomal RNA gene, partial sequence
Euryarchaeota	OS112	JQ 626966	Uncultured archaeon clone KS_226 16S ribosomal RNA gene, partial sequence
	OS122	JQ 626992	Uncultured archaeon clone KS_183 16S ribosomal RNA gene, partial sequence
	OM26	JQ 626833	Uncultured archaeon clone PK_300 16S ribosomal RNA gene, partial sequence
	OD97	JQ 626951	Uncultured archaeon clone PD_177 16S ribosomal RNA gene, partial sequence
MG - I	OM65	JQ 626839	Uncultured archaeon clone SM_289 16S ribosomal RNA gene, partial sequence
	OK108	JQ 626936	Uncultured archaeon clone PK_170 16S ribosomal RNA gene, partial sequence
	OK15	JQ 626840	Uncultured archaeon clone PK_279 16S ribosomal RNA gene, partial sequence
	OM38	JQ 626989	Uncultured archaeon clone SM_182 16S ribosomal RNA gene, partial sequence
	OK14	JX 103593	Uncultured archaeon clone KS_215 16S ribosomal RNA gene, partial sequence
	OM33	JQ 415913	Uncultured archaeon clone SM_300 16S ribosomal RNA gene, partial sequence
	OK104	JQ 626837	Uncultured archaeon clone PK_286 16S ribosomal RNA gene, partial sequence
	OK16	JQ 626860	Uncultured archaeon clone KS_248 16S ribosomal RNA gene, partial sequence
	OM11	JQ 626838	Uncultured archaeon clone PD_219 16S ribosomal RNA gene, partial sequence
	OM50	JQ 626879	Uncultured archaeon clone SM_140R 16S ribosomal RNA gene, partial sequence
	OM7	JQ 626981	Uncultured archaeon clone SM_305R 16S ribosomal RNA gene, partial sequence
	OS123	JQ 626994	Uncultured archaeon clone KS_95 16S ribosomal RNA gene, partial sequence
	OD93	JQ 626841	Uncultured archaeon clone PD_235 16S ribosomal RNA gene, partial sequence
	OD86	JX 103582	Uncultured archaeon clone PD_332R 16S ribosomal RNA gene, partial sequence
	OS127	JQ 626843	Uncultured archaeon clone KS_211 16S ribosomal RNA gene, partial sequence
	OM12	JQ 627002	Uncultured archaeon clone PK_168 16S ribosomal RNA gene, partial sequence
	OS129	JQ 626962	Uncultured archaeon clone KS_160 16S ribosomal RNA gene, partial sequence
	OS13	JX 103594	Uncultured archaeon clone KS_247 16S ribosomal RNA gene, partial sequence
	OK99	JQ 626861	Uncultured archaeon clone PK_287 16S ribosomal RNA gene, partial sequence
	OS110	JQ 626963	Uncultured archaeon clone KS_176 16S ribosomal RNA gene, partial sequence

	OM19	JQ 627006	Uncultured archaeon clone KS_200 16S ribosomal RNA gene, partial sequence
	OM45	JQ 627017	Uncultured archaeon clone SM_166 16S ribosomal RNA gene, partial sequence
	OD84	JX 103588	Uncultured archaeon clone PD_109 16S ribosomal RNA gene, partial sequence
	OS120	JQ 626990	Uncultured archaeon clone KS_124 16S ribosomal RNA gene, partial sequence
	OM18	JQ 626954	Uncultured archaeon clone KS_115 16S ribosomal RNA gene, partial sequence
	OS111	JQ 626965	Uncultured archaeon clone KS_34 16S ribosomal RNA gene, partial sequence
MG - I	OM17	JQ 415912	Uncultured archaeon clone PK_187 16S ribosomal RNA gene, partial sequence
	OM76	JQ 626911	Uncultured archaeon clone SM_187R 16S ribosomal RNA gene, partial sequence
SAGMCG-1	OM40	JQ 627003	Uncultured archaeon clone SM_117 16S ribosomal RNA gene, partial sequence
MCG	OS119	JQ 626976	Uncultured archaeon clone KS_120 16S ribosomal RNA gene, partial sequence
	OK100	JQ 626988	Uncultured archaeon clone PK_96 16S ribosomal RNA gene, partial sequence
	OM67	JQ 626898	Uncultured archaeon clone SM_56 16S ribosomal RNA gene, partial sequence
	OM56	JQ 626886	Uncultured archaeon clone SM_294 16S ribosomal RNA gene, partial sequence
	OM73	JQ 626908	Uncultured archaeon clone SM_235 16S ribosomal RNA gene, partial sequence
	OM8	JQ 626979	Uncultured archaeon clone SM_62 16S ribosomal RNA gene, partial sequence
	OM74	JQ 626909	Uncultured archaeon clone SM_7 16S ribosomal RNA gene, partial sequence
	OM77	JX 103576	Uncultured archaeon clone SM_90 16S ribosomal RNA gene, partial sequence
	OM69	JQ 626904	Uncultured archaeon clone SM_243 16S ribosomal RNA gene, partial sequence
	OS114	JQ 626968	Uncultured archaeon clone KS_149 16S ribosomal RNA gene, partial sequence
	OM39	JQ 626998	Uncultured archaeon clone SM_112 16S ribosomal RNA gene, partial sequence
	OM78	JQ 626854	Uncultured archaeon clone SM_336 16S ribosomal RNA gene, partial sequence
	OM6	JQ 626993	Uncultured archaeon clone KS_245 16S ribosomal RNA gene, partial sequence
	OK101	JQ 626997	Uncultured archaeon clone PK_331 16S ribosomal RNA gene, partial sequence
	OS126	JQ 627010	Uncultured archaeon clone KS_204 16S ribosomal RNA gene, partial sequence
	OS116	JQ 626973	Uncultured archaeon clone KS_154 16S ribosomal RNA gene, partial sequence
	OM34	JQ 626856	Uncultured archaeon clone SM_306 16S ribosomal RNA gene, partial sequence
	OS124	JQ 626995	Uncultured archaeon clone KS_134 16S ribosomal RNA gene, partial sequence

	OK107	JQ 626930	Uncultured archaeon clone PK_238 16S ribosomal RNA gene, partial sequence
	OM37	JQ 626987	Uncultured archaeon clone SM_80 16S ribosomal RNA gene, partial sequence
	OK105	JQ 626925	Uncultured archaeon clone PK_84 16S ribosomal RNA gene, partial sequence
	OK109	JQ 626957	Uncultured archaeon clone PK_215 16S ribosomal RNA gene, partial sequence
	OM51	JQ 626880	Uncultured archaeon clone SM_106 16S ribosomal RNA gene, partial sequence
	OM44	JQ 627015	Uncultured archaeon clone SM_150 16S ribosomal RNA gene, partial sequence
	OS10	JQ 626970	Uncultured archaeon clone KS_222 16S ribosomal RNA gene, partial sequence
	OS118	JQ 626975	Uncultured archaeon clone KS_137 16S ribosomal RNA gene, partial sequence
MCG	OM62	JQ 626894	Uncultured archaeon clone SM_147 16S ribosomal RNA gene, partial sequence
	OM35	JQ 626977	Uncultured archaeon clone SM_125 16S ribosomal RNA gene, partial sequence
	OM72	JQ 626907	Uncultured archaeon clone SM_224 16S ribosomal RNA gene, partial sequence
	OM42	JQ 627009	Uncultured archaeon clone SM_82 16S ribosomal RNA gene, partial sequence
	OS115	JQ 626969	Uncultured archaeon clone KS_164 16S ribosomal RNA gene, partial sequence
	OD81	JQ 626982	Uncultured archaeon clone PD_329 16S ribosomal RNA gene, partial sequence
	OS117	JQ 626974	Uncultured archaeon clone KS_106 16S ribosomal RNA gene, partial sequence
	OM54	JQ 626883	Uncultured archaeon clone SM_254 16S ribosomal RNA gene, partial sequence
	OM64	JQ 626896	Uncultured archaeon clone SM_78R 16S ribosomal RNA gene, partial sequence
	OS121	JQ 626991	Uncultured archaeon clone KS_128 16S ribosomal RNA gene, partial sequence
	OS125	JQ 626999	Uncultured archaeon clone KS_182 16S ribosomal RNA gene, partial sequence
	OM66	JQ 626897	Uncultured archaeon clone SM_71 16S ribosomal RNA gene, partial sequence